Unit Root Tests for Panels in the Presence of Short-run and Long-run Dependencies: Nonlinear IV Approach with Fixed N and Large T [1]

Yoosoon Chang
Department of Economics
Rice University

Wonho Song
Korea Institute for
International Economic Policy

Abstract

An IV approach, using as instruments nonlinear transformations of the lagged levels, is explored to test for unit roots in panels with general dependency and heterogeneity across cross-sectional units. We allow not only for the cross-sectional dependencies of innovations, but also for the presence of cointegration across cross-sectional levels. Unbalanced panels and panels with differing individual short-run dynamics and cross-sectionally related dynamics are also permitted. Panels with such cross-sectional dependencies and heterogeneities appear to be quite commonly observed in practical applications. Yet, none of the currently available tests can be used to test for unit roots in such general panels. We also more carefully formulate the unit root hypotheses in panels. In particular, using order statistics we make it possible to test for and against the presence of unit roots in some of the individual units for a given panel. The individual IV t-ratios, which are the bases of our tests, are asymptotically normally distributed and cross-sectionally independent. Therefore, the critical values of the order statistics as well as the usual average statistic can be easily obtained from simple elementary probability computations. We show via a set of simulations that our tests work well, while other existing tests fail to perform properly. As an illustration, we apply our tests to the panels of real exchange rates, and find no evidence for the purchasing power parity hypothesis, which is in sharp contrast with the previous studies.

This version: January 28, 2005.
JEL Classification: C12, C15, C32, C33.
Key words and phrases: Panels with cross-sectional dependency and heterogeneity, unit root test, cointegration, covariate, nonlinear instrument, order statistics.

[1] We would like to thank the editor and two anonymous referees for many helpful comments, which have greatly improved an earlier version of this paper. Earlier versions of the paper were circulated under the title "Panel unit root tests in the presence of cross-sectional dependency and heterogeneity," and presented at the 2002 Winter Meetings of the Econometric Society at Atlanta, GA, the 10th International Panel Data Conference at The Academy of Science at Berlin, Germany, and at the International Conference on Current Advances and Trends in Nonparametric Statistics at Crete, Greece. We are grateful to H.R. Moon, J.Y. Park, P. Pedroni, B. Perron and seminar participants at University of Maryland, Graduate School of Business at University of Chicago, IZA at University of Bonn, European Central Bank, University of Central Florida for their helpful comments. Chang thanks the NSF for research support under Grant SES-0233940. Make correspondence to: Yoosoon Chang, Department of Economics - MS 22, Rice University, 6100 Main Street, Houston, TX 77005-1892, Tel: 713-348-2796, Fax: 713-348-5278, E-mail: yoosoon@rice.edu.

1. Introduction

Panel unit root tests have been one of the most active research areas for the past several years. This is largely due to the availability of panel data with long time span, and the growing use of cross-country and cross-region data over time to test for many important economic inter-relationships, especially those involving convergencies/divergencies of various economic variables. The notable contributors in theoretical research on the subject include Levin, Lin and Chu (2002), Im, Pesaran and Shin (2003), Maddala and Wu (1999), Choi (2001a) and Chang (2002, 2004). There have been numerous related empirical researches as well. Examples include MacDonald (1996), Oh (1996) and Papell (1997), just to name a few. The papers by Banerjee (1999), Phillips and Moon (2000) and Baltagi and Kao (2000) provide extensive surveys on the recent developments on the testing for unit roots in panels. See also Choi (2001b) and Phillips and Sul (2001) for some related work in this line of research.

In this paper, we consider an IV approach using as instruments nonlinear transformations of the lagged levels. The idea was explored earlier by Chang (2002) to develop the tests that can be used for panels with cross-sectional dependencies of unknown form. Our work extends the approach by Chang (2002) in several important directions. First, we allow for the presence of cointegration across cross-sectional units. It appears that there is a high potential for such possibilities in many panels of practical interests. Yet, none of the existing tests, including those developed by Chang (2002), is not applicable for such panels. Second, our tests rely on the models augmented by cross-sectional dynamics and other covariates. As demonstrated by Hansen (1995) and Chang, Sickles and Song (2001), the inclusion of covariates can dramatically increase the power of the tests. Third, we formulate the panel unit root hypotheses more carefully. In particular, we consider the null and alternative hypotheses that some, not all, of the cross-sectional units have unit roots. Such hypotheses are often more relevant for practical applications.

The presence of cointegration is dealt with simply by using an orthogonal set of functions as instrument generating functions. Chang (2002) considers the IV t-ratios based on the instruments generated by a single function for all cross-sectional units, and shows their asymptotic independence for panels with general cross-sectional dependency. However, as we demonstrate in the paper, the asymptotic independence of the IV t-ratios may be violated in the presence of cointegration across cross-sectional units, which would invalidate the tests by Chang (2002). It is shown in the paper that this difficulty can be resolved if we use the instruments generated by a set of functions that are orthogonal to each other. If a set of orthogonal instrument generating functions are used, the resulting IV t-ratios become asymptotically independent in the presence of cointegration as well as the cross-correlation of innovations.

One of the main motivations to use panels to test for unit roots is to increase the power. An important possibility, however, has been overlooked here, i.e., the possibility of using covariates. The idea of using covariates to test for a unit root was first suggested by Hansen (1995), and its implementation using bootstrap was studied later by Chang, Sickles and Song (2001). They made it clear that there is a huge potential gain in power if covariates are appropriately chosen. Of course, the choice of proper covariates may be difficult in practical applications. In the panel context, however, some of potential covariates to account for the inter-relatedness of cross-sectional dynamics naturally come up front. For instance, we may include the lagged differences of other cross sections to allow for interactions in the short-run dynamics and the linear combinations of the lagged cross-sectional levels in the presence of cointegration.

Obviously, the power increase is not the only reason to test for unit roots in panels. We are often interested in testing for unit roots collectively for cross-sectional units included in a certain panel. In this case, it is necessary to formulate the hypotheses more carefully. In particular, we may want to test for and against the existence of unit roots in not all, but only a fraction of cross-sectional units. Such formulation is, however, more appropriate to investigate important hypotheses such as the purchasing power parity and the growth convergence theories, among many others. The hypotheses can be tested more effectively using order statistics such as maximum and minimum of individual tests. As we show in the paper, the order statistics constructed from individual nonlinear IV t-ratios have limit distributions which are nuisance parameter free and given by simple functionals of the standard normal distribution function. The critical values are thus easily derived from those of the standard normal distribution.

As should have now become obvious, our model is truly general. It allows for the cross-sectional dependency in both the long-run and the short-run. We permit not only the cross-correlation of the innovations and/or cross-sectional dynamics in the short-run, but also the comovements of the stochastic trends in the long-run. Our formulation of the hypotheses is also sharper and makes it possible to test for and against the partial existence of unit roots in the panels. Yet our limit theories are all Gaussian and extremely simple to derive. All this flexibility and simplicity are due to the employment of the nonlinear IV methodology, or more specifically, the asymptotic independence and normality of the individual nonlinear IV t-ratios. All other existing approaches do not offer such generality, assuming either cross-sectional independence that is unacceptable in most applications or a specific form of cross-sectional correlation structure that may be of only limited applicability.

We conduct a set of simulations to evaluate the finite sample performances of our tests. It appears that our tests perform well and are preferred to other existing tests. In particular, our tests perform significantly better than other tests when there are cointegrating relations in the panel. The other tests suffer from severe size distortions in such cases. The performances of our order statistics in small samples are mixed: The minimum statistic performs quite well even with moderate T, whereas the maximum statistic requires a large T for the reliable performance. For the purpose of illustration, we apply our tests to the analysis of the purchasing power parity (PPP) hypothesis. Our tests do not provide any evidence in favor of the PPP hypothesis for real exchange rates, which is quite in contrast to the results in the previous literature. See Chang (2002) and Wu and Wu (2001) for some recent examples. All of the previous results were, however, obtained using the tests that assume either cross-sectional independence or no cointegration.

The rest of the paper is organized as follows. Section 2 specifies the assumptions and provides the background theory. The models and hypotheses are introduced, and some preliminary theories are included. Section 3 defines the test statistics for individual cross-sectional units and for panels and develops their asymptotics. The results from simulations and empirical applications are summarized in Sections 4 and 5, respectively, and the concluding remarks follow in Section 6. The mathematical proofs are collected in Appendix.

2. Assumptions and Background Theory

We consider a panel model generated by

$$y_{it} = \alpha_i y_{i,t-1} + u_{it}, \quad i = 1, \ldots, N; \; t = 1, \ldots, T_i, \tag{1}$$

where (u_{it}) will be specified later. As usual, the index i denotes individual cross-sectional units, such as individuals, households, industries or countries, and the index t denotes time periods. The number of time series observations T_i for each individual i may differ across cross-sectional units. Hence, unbalanced panels are allowed in our model.

2.1 Unit Root Hypotheses

We are interested in testing unit root null hypotheses for the panel given in (1). More precisely, we consider the following sets of hypotheses.

Hypotheses (A) $\mathrm{H}_0 : \alpha_i = 1$ for all i *versus* $\mathrm{H}_1 : \alpha_i < 1$ for all i.

Hypotheses (B) $\mathrm{H}_0 : \alpha_i = 1$ for all i *versus* $\mathrm{H}_1 : \alpha_i < 1$ for some i.

Hypotheses (C) $\mathrm{H}_0 : \alpha_i = 1$ for some i *versus* $\mathrm{H}_1 : \alpha_i < 1$ for all i.

Hypotheses (A) and (B) both include the same null hypothesis, which implies that the unit root is present in all individual units. However, their null hypotheses compete with different alternative hypotheses. It is tested in Hypotheses (A) against the hypothesis that all individual units are stationary, while in Hypotheses (B) the alternative is that there are some stationary individual units. On the contrary, the null hypothesis in Hypotheses (C) holds as long as the unit root exists in at least one individual unit, and is tested against the alternative hypothesis that all individual units are stationary. The alternative hypotheses in both Hypotheses (B) and (C) are the negations of their null hypotheses. This is not the case for Hypotheses (A).

Virtually all the existing literature on panel unit root tests effectively looks at Hypotheses (A). Some recent works, including Im, Pesaran and Shin (2003) and Chang (2002), allow for heterogeneous panels, and formulate the null and alternative hypotheses as in Hypotheses (B). However, strictly speaking, their use of average t-ratios can only be justified for the test of Hypotheses (A). To properly test Hypotheses (B), the minimum, instead of the average, of individual t-ratios might have been used. It is indeed not difficult to see that the tests based on the minimum would dominate those relying on the averages in terms of power for the test of Hypotheses (B). In our simulations, the minimum statistic actually yields much higher power than the average statistic, especially when only a small fraction of individual units are stationary.

Hypotheses (C) have never been considered in the literature, though they seem to be more relevant in many interesting empirical applications such as tests for the purchasing power parity and the growth convergence theories. Note that the rejection of H_0 in favor of H_1 in Hypotheses (C) directly implies that all (y_{it})'s are stationary, and therefore, purchasing power parities or growth convergences hold if we let (y_{it})'s be real exchange rates or differences in growth rates, respectively. No test, however, is available to deal with Hypotheses (C) appropriately. Here we propose to use the maximum of individual t-ratios for the test of Hypotheses (C).

2.2 Short-run Dynamics

We now completely specify the data generating process for our model introduced in (1). The initial values $(y_{10},\ldots,y_{N0})'$ of $(y_{1t},\ldots,y_{Nt})'$ do not affect our subsequent asymptotic analysis as long as they are stochastically bounded, and therefore we set them at zero for expositional brevity. We let $y_t = (y_{1t},\ldots,y_{Nt})'$ and assume that there are $N-M$ cointegrating relationships in the unit root process (y_t), which are represented by the cointegrating vectors (c_j), $j=1,\ldots,N-M$. The usual vector autoregression and error correction representation allow us to specify the short-run dynamics of (y_t) as

$$\triangle y_{it} = \sum_{j=1}^{N}\sum_{k=1}^{P_i} a_{ij}\triangle y_{j,t-k} + \sum_{j=1}^{N-M} b_{ij}c_j' y_{t-1} + \varepsilon_{it} \qquad (2)$$

for each cross-sectional unit, where (ε_{it}) are white noise, $i=1,\ldots,N$, and $\triangle$ is the difference operator.

To ensure our representation in (2), we assume that

Assumption 2.1 Let (y_t) permit a finite order VAR representation and has $N-M$ linearly independent cointegrating relationships. Moreover, if we let $u_t = \triangle y_t$, then we may write $u_t = \Pi(L)\varepsilon_t$, where $\varepsilon_t = (\varepsilon_{1t},\ldots,\varepsilon_{Nt})'$, L is the lag operator, and $\Pi(z) = \sum_{k=0}^{\infty}\Pi_k z^k$ with $\Pi_0 = I$ and $\sum_{k=0}^{\infty} k\|\Pi_k\| < \infty$.

As is well known, we may deduce from the Granger representation theorem that (y_t) can be written as in (2), and that rank $\Pi(1) = M$.

It follows from the Beveridge-Nelson decomposition that

$$u_t = \Pi(1)\varepsilon_t + (\tilde{u}_{t-1} - \tilde{u}_t)\,,$$

where

$$\tilde{u}_t = \tilde{\Pi}(L)\varepsilon_t$$

with $\tilde{\Pi}(z) = \sum_{k=0}^{\infty}\tilde{\Pi}_k z^k$ and $\tilde{\Pi}_k = \sum_{j=k+1}^{\infty}\Pi_j$. Consequently,

$$y_t = \Pi(1)\sum_{k=1}^{t}\varepsilon_k + (\tilde{u}_0 - \tilde{u}_t).$$

Note that $(\tilde{\Pi}_k)$ is absolutely summable due to the 1-summability of (Π_k) given in Assumption 2.1, and therefore, $(\tilde{u}_t)$ is well defined and stationary. Moreover, if $M < N$ and if C is an $N \times (N - M)$ matrix such that $C'\Pi(1) = 0$, then each column (c_j), $j = 1, \ldots, N - M$, of C represents a cointegrating vector for (y_t).

The data generating process for the innovations (ε_t) is assumed to satisfy the following assumption.

Assumption 2.2 (ε_t) is an iid $(0, \Sigma)$ sequence of random variables with $\mathbf{E}|\varepsilon_t|^\ell < \infty$ for some $\ell > 4$, and its distribution is absolutely continuous with respect to Lebesgue measure and has characteristic function φ such that $\lim_{s\to\infty} |s|^r \varphi(s) = 0$, for some $r > 0$.

Assumption 2.2 lays out the technical conditions that are required to invoke the asymptotic theories for the nonstationary nonlinear models developed by Park and Phillips (1999).

Our unit root tests at individual levels will be based on the regression

$$y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{P_i} \alpha_{i,k} \triangle y_{i,t-k} + \sum_{k=1}^{Q_i} \beta'_{i,k} w_{i,t-k} + \varepsilon_{it} \tag{3}$$

for $i = 1, \ldots, N$, where we interpret (w_{it}) as *covariates* added to the augmented Dickey-Fuller (ADF) regression for the i-th cross-sectional unit. It is important to note that the vector autoregression and error correction formulation of the cointegrated unit root panels in (2) suggests that we use such covariates. We may obviously rewrite (1) and (2) as (3) with several lagged differences of other cross sections and linear combinations of the lagged levels of all cross sections as covariates. In the subsequent development of our theory, we will assume that the data generating process is given by (2) under Assumptions 2.1 and 2.2. This, however, is just for the expositional convenience. We may easily accommodate other covariates accounting for idiosyncratic characteristics of cross-sectional units, as long as they satisfy the conditions laid out in Hansen (1995) or Chang, Sickles and Song (2001).

The unit root regression with covariates was first considered in Hansen (1995) and studied subsequently by Chang, Sickles and Song (2001). It was referred to by them as covariates augmented Dickey-Fuller (CADF) regression. Both Hansen (1995) and Chang, Sickles and Song (2001) show that using covariates offers a great potential in power gain for the test of a unit root. In many panels of interest, we naturally expect to have short-run dynamics that are inter-related across different cross-sectional units, which would make it necessary to include the dynamics of others to properly model own dynamics. It is even necessary to take into consideration the long-run trends of other cross-sectional units in the presence of cointegration, since then error correction mechanism comes into play and the stochastic trends of other cross-sectional units would interfere with own short-run dynamics.

2.3 Basic Tools for Asymptotics

Here we introduce some basic theories that are needed to develop the asymptotics of our statistics. Define a stochastic process U_T for u_t as

$$U_T(r) = T^{-1/2} \sum_{t=1}^{[Tr]} u_t$$

on $[0, 1]$, where $[s]$ denotes the largest integer not exceeding s. The process $U_T(r)$ takes values in $D[0, 1]^N$, where $D[0, 1]$ is the space of cadlag functions on $[0, 1]$. Under Assumptions 2.1 and 2.2, an invariance principle holds for U_T, viz.,

$$U_T \to_d U$$

as $T \to \infty$, where U is an N-dimensional vector Brownian motion with covariance matrix Ω given by

$$\Omega = \Pi(1)\Sigma\Pi(1)'.$$

Under Assumption 2.1, the covariance matrix Ω is in general singular with rank M.

Our asymptotic theory involves the local time of Brownian motion, which we will introduce briefly below. The reader is referred to Park and Phillips (1999, 2001), Chang, Park and Phillips (2001), and the references cited there for the concept of local time and its use in the asymptotics for nonlinear models with integrated time series. The local time L_i of U_i, for $i = 1, \ldots, N$, is defined by

$$L_i(t, s) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t 1\{|U_i(r) - s| < \epsilon\}\, dr.$$

Roughly, the local time L_i measures the time that the Brownian motion U_i spends in the neighborhood of s, up to time t. It is well known that L_i is continuous in both t and s. For any local integrable function G on $\mathbf{R}$, we have an important formula

$$\int_0^t G(U_i(r))\, dr = \int_{-\infty}^{\infty} G(s) L_i(t, s)\, ds, \tag{4}$$

which is called the occupation times formula.

2.4 Instrument Generating Functions

We consider the IV estimation of the augmented autoregression (3). To deal with the cross-sectional dependency, we use the instrument generated by a nonlinear function F_i

$$F_i(y_{i,t-1})$$

for the lagged level $y_{i,t-1}$ of each cross-sectional unit $i = 1, \ldots, N$. For the augmented regressors $x'_{it} = (\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-P_i}; w'_{i,t-1}, \ldots, w'_{i,t-Q_i})$, we use the variables themselves as instruments. Hence, for the entire regressors $(y_{i,t-1}, x'_{it})'$, we use the instruments given by

$$(F_i(y_{i,t-1}), x'_{it})',$$

similarly as in Chang (2002).

The transformations (F_i) will be referred to as the *instrument generating functions* (IGF). We assume that

Assumption 2.3 Let (F_i) be regularly integrable and satisfy (a) $\int_{-\infty}^{\infty} xF_i(x)dx \neq 0$ for all i and (b) $\int_{\infty}^{\infty} F_i(x)F_j(x)dx = 0$ for all $i \neq j$.

The class of *regularly integrable* transformations was first introduced in Park and Phillips (1999), to which the reader is referred for details. They are just transformations on $\mathbf{R}$ satisfying some mild technical regularity conditions.

Assumption 2.3 (a) needs to hold, since otherwise we would have *instrument failure* and the resulting IV estimator becomes inconsistent. It is analogous to the non-orthogonality (between the instruments and regressors) requirement for the validity of IV estimation in standard stationary regressions. See Chang (2002) for more detailed discussions. Assumption 2.4 (b) is necessary to allow for the presence of cointegration. If cointegration is present, the procedure in Chang (2002) relying on the same IGF for all cross-sectional units becomes invalid. This will be explained in detail in the next section.

The Hermite functions of odd orders $k = 2i-1$, $i = 1, \ldots, N$, satisfy all the conditions in Assumption 2.3, and therefore, can be used as a proper set of IGF's. The Hermite function G_k of order k, $k = 0, 1, 2, \ldots$, is defined as

$$G_k(x) = (2^k k! \sqrt{\pi})^{-1/2} H_k(x) e^{-x^2/2}, \tag{5}$$

where H_k is the Hermite polynomial of order k given by

$$H_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}.$$

It is well known that the class of Hermite functions introduced above forms an orthonormal basis for $L^2(\mathbf{R})$, i.e., the Hilbert space of square integrable functions on $\mathbf{R}$. We thus have

$$\int_{-\infty}^{\infty} G_j(x)G_k(x)dx = \delta_{jk}$$

for all j and k, where δ_{jk} is the Kronecker delta. Therefore, we may define the IGF's (F_i) by

$$F_i = G_{2i-1}$$

for $i = 1, \ldots, N$.

2.5 Normalization

The orthogonality of the IGF's yields the orthogonality of the IV t-ratios only when (y_{it}) is asymptotically of the same scale across $i = 1, \ldots, N$. Note that the orthogonality between functions is not preserved under arbitrary rescaling of their arguments. This will be seen more clearly in the proof of Lemma 3.2. Roughly, we have for $i = 1, \ldots, N$ and for $r \in [0, 1]$

$$y_{i[T_i r]} = \sqrt{T_i}\,\frac{y_{i[T_i r]}}{\sqrt{T_i}} \approx_d \sqrt{T_i} U_i(r),$$

where U_i is the limit Brownian motion introduced in Section 2.3. Therefore, (y_{it})'s are asymptotically of the same scale if U_i's have the same variance and T_i's are identical for all cross-sectional units $i = 1, \ldots, N$.

The variance of U_i, i.e., the long-run variance of (y_{it}) can be easily estimated consistently. Therefore, we may assume without loss of generality that (y_{it})'s have the same long-run variance for $i = 1, \ldots, N$, since if necessary we may always normalize them using their estimated long-run variances, so that they all have the unit long-run variance. Unless stated otherwise, this convention will be made throughout the paper.

For the unbalanced panels, T_i's are different across $i = 1, \ldots, N$. In this case, we set an arbitrary cross-sectional unit, say, the first unit, to be the scale numeraire, and let

$$y_{it}^* = \frac{\sqrt{T_1}}{\sqrt{T_i}}\, y_{it}. \tag{6}$$

Note that

$$y_{i[T_i r]}^* = \sqrt{T_1}\,\frac{y_{i[T_i r]}}{\sqrt{T_i}} \approx_d \sqrt{T_1} U_i(r).$$

for $i = 1, \ldots, N$ and for $r \in [0, 1]$. Therefore, given our convention that U_i's all have the same variance, (y_{it}^*)'s are asymptotically of the same scale for $i = 1, \ldots, N$ even in unbalanced panels. For expositional brevity, we assume in the subsequent presentation of our theories that the scale adjustment is already done for all (y_{it})'s, and continue to use (y_{it}) in the place of (y_{it}^*) for $i = 1, \ldots, N$. This should cause no confusion.

3. Test Statistics and Their Asymptotics

In this section, we explicitly define test statistics and establish their asymptotic theories. We first look at IV t-ratios for individual cross-sectional units, and derive their asymptotics. We then discuss how one may combine the individual IV t-ratios in formulating tests for the panel unit root hypotheses, specified earlier as Hypotheses (A) – (C), and subsequently develop the asymptotics for the resulting statistics.

3.1 Individual IV t-ratios and Their Asymptotics

We first define individual IV t-ratios explicitly. Let

$$y_i = \begin{pmatrix} y_{i,1} \\ \vdots \\ y_{i,T_i} \end{pmatrix}, \quad y_{\ell i} = \begin{pmatrix} y_{i,0} \\ \vdots \\ y_{i,T_i-1} \end{pmatrix}, \quad X_i = \begin{pmatrix} x_{i,1}' \\ \vdots \\ x_{i,T_i}' \end{pmatrix}, \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i,1} \\ \vdots \\ \varepsilon_{i,T_i} \end{pmatrix},$$

where $x'_{it} = (\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-P_i}; w'_{i,t-1}, \ldots, w'_{i,t-Q_i})$.[2] Then the covariates augmented autoregression (3) can be written in matrix form as

$$y_i = y_{\ell i}\alpha_i + X_i\gamma_i + \varepsilon_i = Y_i\delta_i + \varepsilon_i, \tag{7}$$

where $\gamma_i = (\alpha_{i,1}, \ldots, \alpha_{i,P_i}; \beta'_{i,1}, \ldots, \beta'_{i,Q_i})'$, $Y_i = (y_{\ell i}, X_i)$, and $\delta_i = (\alpha_i, \gamma'_i)'$. For the regression (7), we consider the estimator $\hat{\delta}_i$ of δ_i given by

$$\hat{\delta}_i = \begin{pmatrix} \hat{\alpha}_i \\ \hat{\gamma}_i \end{pmatrix} = (M'_iY_i)^{-1}M'_iy_i = \begin{pmatrix} F_i(y_{\ell i})'y_{\ell i} & F_i(y_{\ell i})'X_i \\ X'_iy_{\ell i} & X'_iX_i \end{pmatrix}^{-1} \begin{pmatrix} F_i(y_{\ell i})'y_i \\ X'_iy_i \end{pmatrix}, \tag{8}$$

where $M_i = (F_i(y_{\ell i}), X_i)$ with $F_i(y_{\ell i}) = (F_i(y_{i,0}), \ldots, F_i(y_{i,T_i-1}))'$. The estimator $\hat{\delta}_i$ is thus defined to be the IV estimator constructed from the instruments M_i.

The IV estimator $\hat{\alpha}_i$ for the AR coefficient α_i corresponds to the first element of $\hat{\delta}_i$ given in (8). Under the null, we have

$$\hat{\alpha}_i - 1 = B_{T_i}^{-1}A_{T_i}, \tag{9}$$

where

$$\begin{aligned} A_{T_i} &= F_i(y_{\ell i})'\varepsilon_i - F_i(y_{\ell i})'X_i(X'_iX_i)^{-1}X'_i\varepsilon_i \\ &= \sum_{t=1}^{T_i} F_i(y_{i,t-1})\varepsilon_{it} - \sum_{t=1}^{T_i} F_i(y_{i,t-1})x'_{it}\left(\sum_{t=1}^{T_i} x_{it}x'_{it}\right)^{-1}\sum_{t=1}^{T_i} x_{it}\varepsilon_{it}, \\ B_{T_i} &= F_i(y_{\ell i})'y_{\ell i} - F_i(y_{\ell i})'X_i(X'_iX_i)^{-1}X'_iy_{\ell i} \\ &= \sum_{t=1}^{T_i} F_i(y_{i,t-1})y_{i,t-1} - \sum_{t=1}^{T_i} F_i(y_{i,t-1})x'_{it}\left(\sum_{t=1}^{T_i} x_{it}x'_{it}\right)^{-1}\sum_{t=1}^{T_i} x_{it}y_{i,t-1}, \end{aligned}$$

and the variance of A_{T_i} is given by

$$\sigma_i^2\mathbf{E}C_{T_i}$$

under Assumption 2.2, where

$$\begin{aligned} C_{T_i} &= F_i(y_{\ell i})'F_i(y_{\ell i}) - F_i(y_{\ell i})'X_i(X'_iX_i)^{-1}X'_iF_i(y_{\ell i}) \\ &= \sum_{t=1}^{T_i} F_i(y_{i,t-1})^2 - \sum_{t=1}^{T_i} F_i(y_{i,t-1})x'_{it}\left(\sum_{t=1}^{T_i} x_{it}x'_{it}\right)^{-1}\sum_{t=1}^{T_i} x_{it}F_i(y_{i,t-1}). \end{aligned}$$

For testing the unit root hypothesis $\alpha_i = 1$ for each $i = 1, \ldots, N$, we construct the t-ratio statistic from the nonlinear IV estimator $\hat{\alpha}_i$ defined in (9). More specifically, we construct such an IV t-ratio for testing for a unit root in (1) or (3) as

$$\tau_i = \frac{\hat{\alpha}_i - 1}{s(\hat{\alpha}_i)}, \tag{10}$$

[2] The regressor x_{it} includes the lagged differenced terms $\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-P_i}$ of y_{it}. By convention, we assume that y_{it} is observed for $t = -P_i, \ldots, T_i$ and set the range of time index to be $t = 1, \ldots, T_i$. This convention will be made throughout the paper.

where $s(\hat{\alpha}_i)$ is the standard error of the IV estimator $\hat{\alpha}_i$ given by

$$s(\hat{\alpha}_i)^2 = \hat{\sigma}_i^2 B_{T_i}^{-2} C_{T_i}. \tag{11}$$

The $\hat{\sigma}_i^2$ is the usual variance estimator given by $T_i^{-1} \sum_{t=1}^{T_i} \hat{\varepsilon}_{it}^2$, where $\hat{\varepsilon}_{it}$ is the fitted residual from the augmented regression (3), viz.,

$$\hat{\varepsilon}_{it} = y_{it} - \hat{\alpha}_i y_{i,t-1} - \sum_{k=1}^{P_i} \hat{\alpha}_{i,k} \triangle y_{i,t-k} - \sum_{k=1}^{Q_i} \hat{\beta}'_{i,k} w_{i,t-k} = y_{it} - \hat{\alpha}_i y_{i,t-1} - x'_{it} \hat{\gamma}_i.$$

It is natural in our context to use the IV estimate $(\hat{\alpha}_i, \hat{\gamma}'_i)'$ given in (8) to get the fitted residual $\hat{\varepsilon}_{it}$. However, we may obviously use any other estimator of $(\alpha_i, \gamma'_i)'$ as long as it yields a consistent estimate for the residual error variance.

The limit null distribution of the IV t-ratio τ_i for testing $\alpha_i = 1$ defined in (10) is derived easily from the asymptotics for nonlinear transformations of integrated processes established in Park and Phillips (1999, 2001) and Chang, Park and Phillips (2001) and is given in

Lemma 3.1 Under Assumptions 2.1-2.3, we have

$$\tau_i \ \to_d \ \mathbf{N}(0,1)$$

as $T_i \to \infty$ for all $i = 1, \ldots, N$.

The normality of the limiting null distribution of the IV t-ratio τ_i is a direct consequence of using the instrument $F_i(y_{i,t-1})$, a regularly integrable transformation of the lagged level $y_{i,t-1}$ which is an integrated process under the unit root null hypothesis. Our limit theory here is thus fundamentally different from the usual unit root asymptotics. This is due to the local time asymptotics and mixed normality of the sample moment $\sum_{t=1}^{T_i} F_i(y_{i,t-1})\varepsilon_{it}$ and the asymptotic orthogonalities between the instrument $F_i(y_{i,t-1})$ and the augmented variables $(\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-P_i}; w'_{i,t-1}, \ldots, w'_{i,t-Q_i})$ which are all stationary. The nonlinearity of the instrument is therefore essential for our Gaussian limit theory. Moreover, the limit standard normal distributions are independent across cross-sectional units $i = 1, \ldots, N$, as we show in the next section.

Our unit root test based on the IV t-ratio statistic is consistent. Under the alternative of stationarity, the IV t-ratio τ_i given in (10) indeed diverges at the $\sqrt{T_i}$-rate. This can be shown using the same argument as the one in Chang (2002, p.270), to which the interested reader is referred. Consequently, the IV t-ratio τ_i diverges at the same rate as the usual OLS-based t-type unit root tests such as the augmented Dickey-Fuller test, under the alternative of stationarity.

3.2 Test Statistics for Panels and Their Asymptotics

For the tests of Hypotheses (A) – (C), we let τ_i be the IV t-ratio for the i-th cross-sectional unit, and define

$$S \ = \ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \tau_i,$$

$$S_{\min} = \min_{1\le i\le N} \tau_i,$$
$$S_{\max} = \max_{1\le i\le N} \tau_i.$$

The average statistic S is proposed for the test of Hypotheses (A), and comparable to other existing tests. The minimum statistic $S_{\min}$ is more appropriate for the test of Hypotheses (B). To test for Hypotheses (B), the average statistic S can also be used, but the test based on $S_{\min}$ would be preferable as discussed earlier. The maximum statistic $S_{\max}$ can be used to test Hypotheses (C). Obviously, the average statistic S and the minimum statistic $S_{\min}$ cannot be used to test for Hypotheses (C), since they would have incorrect sizes.

Let M be $0 \le M \le N$ and define

$$T_{\min} = \min_{1\le i\le N} T_i, \quad T_{\max} = \max_{1\le i\le N} T_i.$$

We assume

Assumption 3.1 Let $\alpha_i = 1$ for $1 \le i \le M$, and set $M = 0$ if $\alpha_i < 1$ for all $1 \le i \le N$.

Assumption 3.2 Assume

$$T_{\min} \to \infty, \quad T_{\max}/T_{\min}^2 \to 0,$$

which will simply be signified by $T \to \infty$ in our subsequent asymptotics.

Assumption 3.1 implies that there are M cross-sectional units having unit roots.[3] Assumption 3.2 gives the premier for our asymptotics. Our asymptotics are based on T-asymptotics and require that the time spans for all cross-sectional units be large for our asymptotics to work. However, we allow for unbalanced panels and they only need to be balanced asymptotically. Our conditions here are fairly weak, and we may therefore expect them to hold widely. The conditions in Assumption 3.2 are stronger than those in Assumption 4.1 of Chang (2002), which require $T_{\min} \to \infty$ and $T_{\max}(\log T_{\max})^4/T_{\min}^3 \to 0$. This is because we allow for the presence of cointegration. Even though it can be effectively dealt with by using a set of orthogonal IGF's, we need slightly more stringent assumption on the balancedness of the underlying panels.

We have

Lemma 3.2 Under Assumptions 2.1–2.3 and 3.1–3.2, the results in Lemma 3.1 hold jointly for all $i = 1, \ldots, M$ and independently across $i = 1, \ldots, M$.

The asymptotic independence of τ_i's is crucial for the subsequent development of our theory. Note that here we allow for the presence of cointegration as well as unknown form of cross-sectional dependencies in the innovations. We now explain the reason why

[3] We defined earlier M to be the number of independent unit roots, net of the number of cointegration relationships, and should not be confused with the usage here. The presence of cointegration no longer affects our asymptotics, due to the orthogonality of the set of IGF's.

we may expect their asymptotic independence even under such general cross-sectional dependencies. Assume for simplicity that the panels are balanced, i.e., $T_i = T$ for all i. As shown in Chang, Park and Phillips (2001), we have

$$\begin{aligned}
\frac{1}{\sqrt[4]{T}}\sum_{t=1}^{T} F_i(y_{i,t-1})\varepsilon_{it} &\approx_d \sqrt[4]{T}\int_0^1 F_i(\sqrt{T}U_i)dV_i, \\
\frac{1}{\sqrt[4]{T}}\sum_{t=1}^{T} F_j(y_{j,t-1})\varepsilon_{jt} &\approx_d \sqrt[4]{T}\int_0^1 F_j(\sqrt{T}U_j)dV_j,
\end{aligned}$$

which become independent if and only if their quadratic covariation

$$\sigma_{ij}\sqrt{T}\int_0^1 F_i(\sqrt{T}U_i(r))F_j(\sqrt{T}U_j(r))dr \to_{a.s.} 0 \tag{12}$$

as $T \to \infty$, where σ_{ij} denotes the covariance between V_i and V_j representing the limit Brownian motions of (ε_{it}) and (ε_{jt}), respectively.

It is indeed well known that

$$\int_0^1 F_i(\sqrt{T}U_i(r))F_j(\sqrt{T}U_j(r))dr = O_p(\log T/T) \quad \text{a.s.} \tag{13}$$

for any Brownian motions U_i and U_j so long as they are not degenerate, and this implies that the condition (12) holds even when $\sigma_{ij} \neq 0$. Chang (2002) uses this result to develop the unit root tests for panels with cross-sectionally correlated innovations. However, (13) does not hold in the presence of cointegration between (y_{it}) and (y_{jt}). In this case, their limiting Brownian motions U_i and U_j become degenerate. If the cointegrating relationship is given by the unit coefficient, for instance, then we would have $U_i = U_j$, and therefore,

$$\begin{aligned}
\sqrt{T}\int_0^1 F_i(\sqrt{T}U_i(r))F_j(\sqrt{T}U_j(r))dr &= \sqrt{T}\int_0^1 (F_iF_j)(\sqrt{T}U_i(r))dr \\
&= \sqrt{T}\int_{-\infty}^{\infty} (F_iF_j)(\sqrt{T}s)L_i(1,s)ds \\
&= \int_{-\infty}^{\infty} (F_iF_j)(s)L_i(1,s/\sqrt{T})ds \\
&= \left(\int_{-\infty}^{\infty} (F_iF_j)(s)ds\right) L_i(1,0) + o_{a.s.}(1)
\end{aligned}$$

by the occupation times formula (4), change of variables and the continuity of $L(1,\cdot)$. The asymptotic independence of τ_i and τ_j generally breaks down, and holds only when F_i and F_j are orthogonal. This is the reason why the method by Chang (2002) becomes invalid in the presence of cointegration. We use an orthogonal set of IGF's to preserve the asymptotic independence here.

The asymptotic theories for the statistics S, $S_{\min}$ and $S_{\max}$ may be easily derived from Lemma 3.2. We now let Φ be the distribution function for the standard normal distribution, and let λ be the size of the tests. For a given size λ, we define $c_{\max}^M(\lambda)$ and $c_{\min}(\lambda)$ by

$$\Phi(c_{\max}^M(\lambda))^M = \lambda, \quad (1 - \Phi(c_{\min}(\lambda)))^N = 1 - \lambda.$$

These provide the critical values of the statistics $S_{\min}$ and $S_{\max}$ for the tests of Hypotheses (B) and (C). The critical values $c(\lambda)$ of the average test S for Hypotheses (A) are defined as usual from $\Phi(c(\lambda)) = \lambda$. The following table shows the tests and critical values that should be used to test each of Hypotheses (A) – (C).

Hypotheses	Test Statistics	Critical Values
Hypotheses (A)	S	$c(\lambda)$
Hypotheses (B)	$S_{\min}$	$c_{\min}(\lambda)$
Hypotheses (C)	$S_{\max}$	$c_{\max}^M(\lambda)$

The critical values $c_{\max}^M(\lambda)$ and $c_{\min}(\lambda)$ for sizes $\lambda = 1\%, 5\%$ and 10% are tabulated in Table 1 for some cases of M, N up to 100.

The following lemma summarizes the asymptotic behaviors of $S, S_{\min}$ and $S_{\max}$.

Theorem 3.3 Let Assumptions 2.1–2.3 and 3.1–3.2 hold. If $M = N$,

$$\begin{aligned} \lim_{T\to\infty} \mathbf{P}\{S \le c(\lambda)\} &= \lambda, \\ \lim_{T\to\infty} \mathbf{P}\{S_{\min} \le c_{\min}(\lambda)\} &= \lambda. \end{aligned}$$

If $1 \le M \le N$, then

$$\lim_{T\to\infty} \mathbf{P}\{S_{\max} \le c_{\max}^M(\lambda)\} = \lambda, \quad \lim_{T\to\infty} \mathbf{P}\{S_{\max} \le c_{\max}^1(\lambda) = c(\lambda)\} \le \lambda.$$

On the other hand, $S, S_{\max} \to_p -\infty$ if $M = 0$, and $S_{\min} \to_p -\infty$ if $M < N$.

Theorem 3.3 implies that all our tests have the prescribed asymptotic sizes. The tests using statistics S and $S_{\min}$ with critical values $c(\lambda)$ and $c_{\min}(\lambda)$, respectively, have the exact size λ asymptotically under the null hypotheses in Hypotheses (A) and (B). However, the null hypothesis in Hypotheses (C) is composite, and the rejection probabilities of the test relying on $S_{\max}$ with critical values $c(\lambda)$ may not be exactly λ even asymptotically. The size λ in this case is the maximum rejection probabilities that may result in under the null hypothesis. Theorem 3.3 also shows that all our tests are consistent for Hypotheses (A) – (C).

3.3 Models with Deterministic Components

The models with deterministic components can be analyzed similarly using properly demeaned or detrended data. As argued and demonstrated in Chang (2002), a proper demeaning or detrending scheme required here must be able to remove the nonzero mean and time trend successfully, while preserving the predictability of the instruments and ultimately the Gaussian limit theory for the nonlinear IV unit root tests. We now introduce our demeaning and detrending schemes. The methods are basically similar to those given in Chang (2002), but with additional attention paid to the covariates.

If the time series (z_{it}) with a nonzero mean is given by

$$z_{it} = \mu_i + y_{it}, \tag{14}$$

where the stochastic component (y_{it}) is generated as in (1), then we may test for the presence of a unit root in (y_{it}) from the covariates augmented regression (3) defined with the demeaned series y_{it}^{μ}, $y_{i,t-1}^{\mu}$, $\triangle y_{i,t-k}^{\mu}$ and $w_{i,t-k}^{\mu}$ of z_{it}, $z_{i,t-1}$, $\triangle z_{i,t-k}$ and $w_{i,t-k}$, viz.,

$$y_{it}^{\mu} = \alpha_i y_{i,t-1}^{\mu} + \sum_{k=1}^{P_i} \alpha_{i,k} \triangle y_{i,t-k}^{\mu} + \sum_{k=1}^{Q_i} \beta_{i,k}' w_{i,t-k}^{\mu} + e_{it}. \tag{15}$$

As in Chang (2002), the demeaned series y_{it}^{μ} and $y_{i,t-1}^{\mu}$ are constructed by subtracting the mean of the partial sample up to time $(t-1)$, i.e., $(t-1)^{-1} \sum_{k=1}^{t-1} z_{ik}$, which is the least squares estimator of μ_i in (14). The $(t-1)$-adaptive demeaning is used to maintain the martingale property and thus the Gaussian limit theory of our nonlinear IV t-ratios. The terms $\triangle y_{i,t-k}^{\mu}$ are simply the differences of the original data, i.e., $\triangle z_{i,t-k}$, for $k = 1, 2, \ldots, P_i$, and the (e_{it}) are regression errors. We now define the demeaned covariate series $w_{i,t-k}^{\mu}$. In order to demean the covariates $w_{i,t-k}$ properly, we first need to know what types of covariates are used. For the given panel, there are three natural groups of candidates for the covariates: (i) the lagged differences $\triangle y_{j,t-k}$ for the I(1) y_{jt}'s not cointegrated with y_{it}, (ii) the lagged cointegration errors, say $(y_{i,t-1} - \phi_i y_{j,t-1})$ for the I(1) y_{jt}'s cointegrated with y_{it},[4] and (iii) the level values y_{jt} for the stationary y_{jt}'s. Then, the following demeaned covariates may be used for each type of covariates, respectively, viz.,

$$\triangle y_{j,t-k} \quad : \quad \triangle z_{j,t-k}, \qquad \text{for } k = 1, 2, \ldots, Q_i, \tag{16}$$

$$(y_{i,t-1} - \phi_i y_{j,t-1}) \quad : \quad \left(z_{i,t-1} - \frac{1}{T_i} \sum_{k=1}^{T_i} z_{ik} \right) - \phi_i \left(z_{j,t-1} - \frac{1}{T_j} \sum_{k=1}^{T_j} z_{jk} \right), \tag{17}$$

$$y_{jt} \quad : \quad z_{jt} - \frac{1}{T_j} \sum_{k=1}^{T_j} z_{jk}. \tag{18}$$

The term $T_j^{-1} \sum_{k=1}^{T_j} z_{jk}$ appearing in (17) and (18) is the grand sample mean of z_{jt}, and it is used to remove the nonzero mean of $z_{j,t-1}$ and z_{jt}.

[4]Here, we assume that the true cointegrating relations are known. In practice, however, the cointegrating vectors have to be estimated.

We may then construct the nonlinear IV t-ratio statistic τ_i^μ for the process (z_{it}) in (14) with a nonzero mean based on the nonlinear IV estimator for α_i computed from the covariates augmented regression (15), just as in (10).

Similarly as in the models with nonzero means, we may also test for unit roots in the models with deterministic time trends using the nonlinear IV unit root test constructed from the properly detrended data. More explicitly, for the time series with a linear time trend

$$z_{it} = \mu_i + \delta_i t + y_{it} \tag{19}$$

where (y_{it}) is generated as in (1), we may test for a unit root in (y_{it}) from the regression (3) defined with the properly detrended series y_{it}^τ, $y_{i,t-1}^\tau$, $\triangle y_{i,t-k}^\tau$ and $w_{i,t-k}^\tau$ of the given data z_{it}, $z_{i,t-1}$, $\triangle z_{i,t-k}$ and $w_{i,t-k}$, viz.,

$$y_{it}^\tau = \alpha_i y_{i,t-1}^\tau + \sum_{k=1}^{P_i} \alpha_{i,k} \triangle y_{i,t-k}^\tau + \sum_{k=1}^{Q_i} \beta_{i,k}' w_{i,t-k}^\tau + e_{it}. \tag{20}$$

The detrended data y_{it}^τ and $y_{i,t-1}^\tau$ are constructed following the adaptive detrending scheme introduced in Chang (2002) which uses the least squares estimators of the drift and trend coefficients, μ_i and δ_i, from the model (19) using again the observations up to time $(t-1)$ only. The adaptive detrending preserves the predictability of our instrument $F(y_{i,t-1}^\tau)$. The lagged differences are also detrended just as in Chang (2002), viz., $\triangle y_{i,t-k}^\tau = \triangle z_{i,t-k} - z_{iT_i}/T_i$, where the grand sample mean of $\triangle z_{it}$, i.e., $T_i^{-1} \sum_{k=1}^{T_i} \triangle z_{ik}$, is used to eliminate the nonzero mean of $\triangle z_{i,t-k}$, for $k = 1, \ldots, P_i$. The detrended covariates $w_{i,t-k}^\tau$ for the model (19) with linear trend are constructed accordingly for each type of the covariates mentioned above (16) as

$$\triangle y_{j,t-k} \quad : \quad \triangle z_{j,t-k} - \frac{1}{T_j} z_{jT_j}, \quad \text{for } k = 1, 2, \ldots, Q_i, \tag{21}$$

$$(y_{i,t-1} - \phi_i y_{j,t-1}) \quad : \quad (z_{i,t-1} - \hat{\mu}_i^{T_i} - \hat{\delta}_i^{T_i}(t-1)) - \phi_i(z_{j,t-1} - \hat{\mu}_j^{T_j} - \hat{\delta}_j^{T_j}(t-1)) \tag{22}$$

$$y_{jt} \quad : \quad z_{jt} - \hat{\mu}_j^{T_j} - \hat{\delta}_j^{T_j} t. \tag{23}$$

The parameters $\hat{\mu}_s^{T_s}$ and $\hat{\delta}_s^{T_s}$ for $s = i,\ j\ (\neq i)$, in (22) and (23) are estimated using the full sample from the model (19), and in (21) the grand sample mean z_{jT_i}/T_i of $\triangle z_{jt}$ is used to eliminate the nonzero mean of $\triangle z_{j,t-k}$, for $k = 1, \ldots, Q_i$.

The nonlinear IV t-ratio τ_i^τ for testing unit roots in the model (19) with linear time trend is then defined as in (10) using the nonlinear IV estimator for α_i computed from the regression (20) based on the adaptively detrended data.

With the adaptive demeaning or adaptive detrending, the predictability of our nonlinear instrument $F_i(y_{i,t-1}^\mu)$ or $F_i(y_{i,t-1}^\tau)$ is retained, and consequently our previous results continue to apply, including the normal distribution theory for the IV t-ratio statistic. We may now derive the limit theories of the statistics τ_i^μ and τ_i^τ for the models with nonzero means and deterministic trends in the similar manner as we did to establish the limit theory given in Lemma 3.1.

Corollary 3.4 The results in Lemmas 3.1 and 3.2 hold also for both τ_i^μ and τ_i^τ.

It follows that the standard normal limit theory of the covariates augmented nonlinear IV t-ratio statistics continues to hold for the models with deterministic components.

4. Simulations

In this section, we conduct a set of simulations to investigate the finite sample performances of the newly proposed panel unit root tests for testing the three hypotheses formulated in Section 2. For the simulations, we consider the model (14) with a nonzero mean and the stochastic component (y_{it}) specified as in (1) with the innovations (u_{it}) generated by the following three DGP's:

$$\begin{aligned}
\text{DGP1} &: \quad u_{it} = \beta_i u_{i,t-1} + \eta_{it}, \\
\text{DGP2} &: \quad u_{it} = \beta_i u_{i,t-1} + \nu_i \xi_t + \eta_{it}, \\
\text{DGP3} &: \quad u_{it} = \beta_i u_{i,t-1} + \nu_i \xi_t + \triangle \eta_{it},
\end{aligned}$$

for $i = 1, \ldots, N$; $t = 1, \ldots, T$, where ξ_t is the scalar common stochastic trend and $\eta_t = (\eta_{1t}, \ldots, \eta_{Nt})'$ an N-dimensional innovation vector with cross-sectional dependence. Note that DGP1 is the same model analyzed in Chang (2002), DGP2 is a version of the widely used dynamic factor model, and DGP3 is new and introduced here to allow for cointegration across cross-sectional units. By exploring these distinct forms of cross-correlations, we aim to see how our tests perform relative to the existing tests in each situation.

DGP1 generates cross-sectional correlations from dependent innovations (η_{it}) with the covariance matrix, say V, which is unrestricted except for being symmetric and nonsingular. The innovations in DGP2 and DGP3 also have the same error covariance V. However, DGP2 and DGP3 have another level of cross-correlations coming from the presence of the common stochastic trend (ξ_t). Using this common factor, DGP2 and DGP3 can generate stronger cross-sectional dependencies compared to those generated by DGP1. In DGP2, the generated series (y_{it}) contain both nonstationary common factors and nonstationary individual errors under the unit root null hypothesis, hence there is no cointegrating relationship among cross sections. In DGP3, however, cross-sectional cointegration is present, and they are generated by the nonstationary common stochastic trend coupled with stationary individual errors. Thus, there exists a cointegrating relationship between any pair of (y_{it}) and (y_{jt}), with $(N-1)$ linearly independent cointegrating relations among N individual units.

The parameters in our DGP's are generated as follows. The AR coefficient β_i is drawn randomly from Uniform$[0.2, 0.4]$. The parameter ν_i, so called factor loadings, that controls the relative importance of common versus idiosyncratic shocks is also drawn randomly from Uniform$[0.5, 3]$. The processes (ξ_t) and (η_t) are independent and drawn from iid $\mathbf{N}(0,1)$ and iid $\mathbf{N}(0,V)$, respectively. The parameters of the $(N \times N)$ covariance matrix V of the innovations (η_t) are also drawn randomly. To ensure that V is a symmetric positive definite matrix and to avoid the near-singularity problem, we generate V following the steps outlined in Chang (2002). The steps are presented here for convenience:

(1) Generate an $(N \times N)$ matrix M from Uniform[0,1].
(2) Construct from M an orthogonal matrix $H = M(M'M)^{-1/2}$.
(3) Generate a set of N eigenvalues, $\lambda_1, \ldots, \lambda_N$. Let $\lambda_1 = r > 0$ and $\lambda_N = 1$ and draw $\lambda_2, \ldots, \lambda_{N-1}$ from Uniform[r,1].
(4) Form a diagonal matrix Λ with $(\lambda_1, \ldots, \lambda_N)$ on the diagonal.
(5) Construct the covariance matrix V using the spectral representation $V = H\Lambda H'$.

Constructed as such, the covariance matrix V will be symmetric and nonsingular with eigenvalues ranging from r to 1. The ratio r of the minimum eigenvalue to the maximum provides a measure for the degree of correlations and heterogeneity in the error covariance matrix V. The covariance matrix V becomes singular as r tends to zero and becomes spherical as r approaches 1. For the simulations, we set $r = 0.1$ as in Chang (2002), and use the correlation matrix obtained from the covariance matrix V in the usual manner.

The panels with the cross-sectional dimensions $N = 10, 20, 50$ and the time series dimensions $T = 100, 200$ are considered for the 5% nominal test size. We set the AR coefficient α_i in (1) at $\alpha_i = 1$ for the cross sections that have unit root, and generate α_i randomly from Uniform[0.8,1] for the stationary cross sections. Since we are using randomly drawn parameter values, we simulate 10 times and report the average of the finite sample performances of the tests. Each simulation run is carried out with 3,000 iterations. We assume that there exist nonzero means in the data, and thus we use the adaptively demeaned series as in (16)-(18) for our nonlinear IV tests.

In our simulations, we consider the following tests:

Nonlinear IV Unit Root Tests	
$S^C, S^C_{\min}, S^C_{\max}$	ave, min and max tests with single IGF and no covariate
$S^F, S^F_{\min}, S^F_{\max}$	ave, min and max tests with single IGF and covariate
$S^H, S^H_{\min}$	ave and min tests with orthogonal IGF's and no covariate
$S^A, S^A_{\min}$	ave and min tests with orthogonal IGF's and covariate
Other Existing Tests	
IPS	test by Im, Pesaran and Shin
MP	test by Moon and Perron

The test S^C, based on a single integrable IGF, is developed in Chang (2002) for Hypotheses (A). It is considered here for the comparison with other IV tests. The tests $S^C_{\min}$ and $S^C_{\max}$ are, respectively, the minimum and maximum counterparts of the average test S^C for Hypotheses (B) and (C) relying on a single integrable IGF. The tests $S^F, S^F_{\min}$ and $S^F_{\max}$ are the ones with covariate, each corresponding to $S^C, S^C_{\min}$ and $S^C_{\max}$. We may compare these two sets of the tests to analyze the effect of including a covariate. The tests $S^C, S^C_{\min}, S^C_{\max}$

and $S^F, S^F_{\min}, S^F_{\max}$ are valid only for DGP1 and DGP2. The presence of cross-sectional cointegration in DGP3 invalidates these tests. Considered subsequently are the tests S^H and S^A, which are the average tests based on the orthogonal IGF's, respectively without and with covariate. The tests $S^H_{\min}$ and $S^A_{\min}$ are their minimum counterparts. These tests are applicable for all DGP's. They are, however, intended to effectively deal with the presence of cross-sectional cointegration in DGP3. The maximum tests with the orthogonal IGF's are not examined. The cross-sectional cointegration presumes the presence of unit roots, and therefore, Hypotheses (C) are not particularly interesting in this situation.

We also consider two other existing tests IPS and MP, by Im, Pesaran and Shin (2003) and Moon and Perron (2001), respectively. The IPS test is based on the average of the individual t-ratios computed from the usual sample ADF regressions with mean and variance modifications. More explicitly, the IPS test is defined as

$$\text{IPS} = \frac{\sqrt{N}(\bar{t}_N - N^{-1}\sum_{i=1}^N \mathbf{E}(t_i))}{\sqrt{N^{-1}\sum_{i=1}^N \text{var}(t_i)}}$$

where t_i is the t-statistic for testing $\alpha_i = 1$ for the i-th sample ADF regression, and $\bar{t}_N = N^{-1}\sum_{i=1}^N t_i$. The values of the expectation and variance, $\mathbf{E}(t_i)$ and $\text{var}(t_i)$, for each individual t_i depend on T_i and the lag order P_i, and are computed via simulations from independent normal samples. See Table 3 in Im, Pesaran, and Shin (2003). The IPS test assumes cross-sectional independence and hence it is not valid under our DGP's.

The MP test models the cross-sectional dependence using an approximate dynamic linear factor model. They model the error process $\triangle y_{it} = u_{it}$ as $u_{it} = \delta_i'\xi_t + \varepsilon_{it}$, where (ε_{it}) are cross-sectionally independent idiosyncratic shocks, δ_i factor loadings for the i-th unit, and ξ_t an unknown number of unobservable dynamic factors that are common to all individual units. Then, it follows under the null of $\alpha_i = 1$ for each $i = 1, \ldots, N$,

$$y_{it} = y_{i,t-1} + u_{it} = y_{i0} + \delta_i' \sum_{k=1}^t \xi_k + \sum_{k=1}^t \varepsilon_{ik}.$$

Under this setup, first the data is demeaned, and the cross-correlations generated by the nonstationary common factors $\sum_{k=1}^t \xi_k$ are removed by projecting the panel data to the space orthogonal to the factor loadings δ_i. Then, they calculate the t-statistic to test for unit roots existing in such 'de-factored' panel data, say $\widetilde{y}$, based on the modified pooled OLS estimator $\widetilde{\alpha}_{pool}$ of α obtained from the pooled regression

$$\widetilde{y} = \alpha\widetilde{y}_{-1} + \widetilde{\varepsilon}, \tag{24}$$

where $\widetilde{y} = (\widetilde{y}_1', \ldots, \widetilde{y}_N')'$ and $\widetilde{\varepsilon} = (\widetilde{\varepsilon}_1', \ldots, \widetilde{\varepsilon}_N')'$.[5] More explicitly, the MP test is defined as

$$\text{MP} = \frac{\sqrt{N}T(\widetilde{\alpha}_{pool} - 1)}{\sqrt{3\widehat{\phi}_\varepsilon^4/\widehat{\omega}_\varepsilon^4}},$$

[5]Note that in (24) the regression errors are now cross-sectionally independent since the de-factored data, $\widetilde{y}$, used in the estimation, are free of the common factors that generate the correlations across cross sections.

where $\widehat{\omega}_\varepsilon^2 = \frac{1}{N}\sum_{i=1}^N \widehat{\omega}_{\varepsilon,i}^2$, $\widehat{\phi}_\varepsilon^4 = \frac{1}{N}\sum_{i=1}^N \widehat{\omega}_{\varepsilon,i}^4$ and $\widehat{\omega}_{\varepsilon,i}^2$ are the long-run variance of (ε_{it}).[6] Here, we note that the MP test may be used to test for unit roots in the models generated by DGP1 or DGP2 only when η_{it}'s are independent across i. It is invalid under DGP3 even with the cross-sectional independence of η_{it}'s.

As the covariate, we use the estimated common factor given by the average of the fitted residuals $\hat{v}_{it}$ from the regression of $\triangle y_{it}$ on $\triangle y_{i,t-1}$.[7] The tests with the covariate are thus expected to perform better for DGP2, compared to the tests constructed without the covariate. If we set $v_{it} = \nu_i \xi_i + \eta_{it}$ under DGP2, we may indeed deduce that

$$\frac{1}{N}\sum_{i=1}^N v_{it} = \bar{\nu}\xi_t + \frac{1}{N}\sum_{i=1}^N \eta_{it},$$

where $\bar{\nu} = N^{-1}\sum_{i=1}^N \nu_i$. Therefore, our covariate more precisely estimates the common factor as N gets large and the realized values of ν_i's are concentrated around $\bar{\nu}$. For the actual implementation of our tests with the covariate, we rank the estimated error variances of v_{it}, which amount to be $\nu_i^2 + \sigma_i^2$, and choose a sub-group of the cross-sectional units among which the estimated variances vary least.[8] This allows us to select the units with most homogeneous factor loadings so long as the σ_i^2 are the same across i. For the selection of the sub-group, we simply choose in our simulations the units corresponding to the middle 40% of the error variances.[9] The factor estimate obtained this manner appears to work reasonably well across all T's and N's that we consider in the simulations.

For the construction of the nonlinear tests S^C and S^F with single IGF, we use the same IGF suggested in Chang (2002). For the tests S^H and S^A with orthogonal IGF's, we use a set of the N-Hermite functions defined in (5). We normalize the data as suggested in Section 2.5 by dividing y_{it}'s, $i = 1, \ldots, N$, by their estimated long-run variances. We then scale the data by multiplying the scale constant c chosen by the following scheme:

$$c = K \times T^{-1/2} \tag{25}$$

for all $i = 1, \ldots, N$.[10] For the tests $S^C, S^C_{\min}, S^C_{\max}, S^F, S^F_{\min}$ and $S^F_{\max}$ with single IGF, we set the constant at $K = 4$. On the other hand, we use the constant $K = 3$ and 2 respectively for the tests S^H and S^A, and $K = 1.5$ and 1 respectively for the tests $S^H_{\min}$ and $S^A_{\min}$. Note that we use the smaller values of K for the tests with orthogonal IGF's, when the covariate is augmented. This is because in this case we are doubly controlling the cross-correlations by using the orthogonal IGF's and by including the estimated factor as a covariate. The estimated factor would pick up substantial amount of the cross-correlations, and thus the

[6]The maximum number of factors for the MP test is set at 8 as in Moon and Perron (2001).

[7]There exist other ways to estimate the common factor in this case, such as those suggested in Bai and Ng (2002), Moon and Perron (2001) and Phillips and Sul (2001). We chose the current method, however, because it is very simple to implement and provides reasonably good performances.

[8]The factor estimate based on a subset of the cross-sectional units appears to work generally better than those obtained using the full sample.

[9]The 40% rule is chosen to provide the best overall size and power performances. The results are, however, not very sensitive to the size of the sub-group.

[10]This scaling scheme is essentially the same as the one used in Chang (2002). Here we set c as given only for small T. As T gets large, the choice of c becomes unimportant.

IGF's need not be as integrable as those used for the tests without the covariate. The cross-correlations are controlled only by the orthogonal IGF's for the tests without the covariate.

The simulation results are reported in Tables 2–4. We first discuss the performances of the nonlinear IV average tests along with IPS and MP, and subsequently present the results for the nonlinear IV minimum and maximum tests. As expected the IPS and MP tests fail in all DGP's we consider here, as can be seen from the considerable size distortions. Their size distortions are large especially in DGP3 with the strongest cross-correlations. The distortions get larger as N increases for a given T and they do not seem to be attenuated as T increases. The MP test seems to have better size property as T gets large, although they are still not very satisfactory. Among the three DGP's, the IPS suffer least from the size distortions in DGP1, while the MP test performs best in DGP2. Also as expected, the simple IV test S^C performs well for DGP1, but it fails in DGP2 and DGP3. The simple IV test S^F augmented with the covariate works well in DGP2 as well as in DGP1. The use of the covariate in S^F seems to improve significantly the performance of the test S^C without covariate, as can be seen clearly from comparing the sizes of S^C and S^F in DGP2. However, S^F also fails in DGP3 with cross-sectional cointegration with large size distortions. This is well expected. Therefore, all the existing tests, including S^F, have severe inferential problems in the presence of the cross-sectional cointegration.

As expected, our new nonlinear IV tests, S^H and S^A, based on the orthogonal IGF's perform reasonably well in DGP3. The S^H test has good sizes for all T's and N's considered. On the other hand, the covariate augmented test S^A has good sizes for the modest size N, but starts to under-reject for the larger N. In DGP1 and DGP2, both tests suffer from downward size distortions, and the problem is worse for the larger N. Their size problem, however, tends to improve as T gets large, and the sizes of S^H and S^A do seem to become reasonably good for the modest size N and the larger T even in DGP1 and DGP2. Note also that the use of orthogonal IGF's is not necessary for DGP1 and DGP2. As mentioned above, the size performance of the covariate augmented test S^A may not be as good as that of the test S^H, indicating that the use of the covariate may deteriorate the size property of our test. However, S^A test performs noticeably better than S^H in terms of power, so there is a substantial gain in using the covariate.[11] From our simulation results, it seems that the covariate augmented simple IV test S^F works best in terms of both sizes and powers when there is no cross-sectional cointegration but exists severe as well as mild cross-correlations. Hence one may use S^F if there is no suspicion about potential cross-sectional cointegration.

The results on the finite sample performances of the minimum tests are reported also in Tables 2–4 along with those of their average counterparts we just discussed above.

[11] In the simulations, due to the computing time, we just used the estimated factor as the covariate, which may not necessarily be the best choice. Though we do not report the details here, we have observed that we may improve the performance of S^A by selecting the best set of covariates from a pool of all potential covariates described in Section 2.2. This can be done by choosing the ones that have the highest correlations with the error process. For more discussions on this, see Hansen (1995) and Chang, Sickles, and Song (2001). This ensures that the effective error has smallest variance, which will in turn lead to the largest power gains from the inclusion of those covariates.

The minimum tests are expected to perform better under the alternative hypothesis that only some fractions of the panel are stationary. To specify such alternatives, we set I_0, the number of I(0) series in the panel, to be 10%, 20%, 50% and 100% of the given N. Overall, all of the four minimum tests, $S^C_{\min}$, $S^F_{\min}$, $S^H_{\min}$ and $S^A_{\min}$, work reasonably well in all DGP's with stable sizes and large discriminatory powers compared to their average test counterparts. The $S^C_{\min}$ test works very well for DGP1. The finite sample sizes of $S^C_{\min}$ are indeed very close to the nominal test size for all DGP's and for all combinations of N and T, even though the test is valid only for DGP1 and DGP2, and not for DGP3. The $S^F_{\min}$ test performs well for DGP1 and DGP2, but it under-rejects for DGP3. The $S^H_{\min}$ and $S^A_{\min}$ tests tend to under-reject for large N for all DGP's considered. Notice that when only a fraction, not all, of the panel are stationary, the powers of the minimum tests are significantly larger than their average counterparts. This is as expected, since the minimum would obviously provide more discriminatory power than the average against such alternatives where only some cross sections are stationary.

Finally, we discuss the performances of the maximum tests which are constructed for Hypotheses (C) with the composite null hypothesis where only some fractions of the panel are nonstationary. For the null hypotheses, we set I_1, the number of I(1) series under the panel, to be 10%, 20%, 50%, and 100% of given N, as in the formulations for the alternatives of Hypotheses (B). The simulation results on the sizes and powers of the maximum tests, $S^C_{\max}$ and $S^F_{\max}$, are provided in Table 5. The test $S^C_{\max}$ yields reasonable sizes and powers for DGP1. In DGP1, $S^C_{\max}$ has stable sizes that are quite close to the nominal test size in most of the N and T combinations we consider here, except when I_1 is very small. The finite sample powers of $S^C_{\max}$ are also quite good in DGP1. Not surprisingly, the test $S^C_{\max}$ does not perform well in DGP2 with strong cross-correlations driven by the common factor. In DGP2, the covariate augmented test $S^F_{\max}$ performs a little better than $S^C_{\max}$. The overall performance of $S^F_{\max}$ is, however, not satisfactory. The test has stable sizes when $I_1 = N$, but severely under-rejects in other cases. It also shows very poor power performance in DGP2. In DGP1, it performs very well like the test $S^C_{\max}$ in terms of both sizes and powers. Our simulation results therefore seem to indicate that the max tests $S^C_{\max}$ and $S^F_{\max}$ are reliable only for the panels with mild cross-correlations such as those generated by DGP1. In this sense, the usefulness of the maximum tests seems somewhat limited.

5. Empirical Illustrations

In this section, we illustrate the usefulness of our new panel unit root tests by applying them to the long standing empirical problem of testing for the purchasing power parity (PPP) hypothesis. Although considerable amount of intellectual efforts have been put out to investigate this problem, it is widely agreed that the question of whether or not the PPP holds has not been settled yet. We revisit this PPP problem with our new tests to examine whether the PPP holds for the post-1973 period of floating exchange rates. As noted by Papell (2000), the data covering only the recent float have some advantages over the long horizon data. For instance, it does not mix observations from different nominal exchange rate regimes and the data are available for more countries. The data we use are the quarterly and monthly end-of-period real exchange rates for twenty industrialized countries,[12] obtained from the International Monetary Fund's International Financial Statistics. For monthly real exchange rates, we consider seventeen out of twenty countries since no monthly data is available for Australia, Ireland, and New Zealand. Our data cover the entire recent float period, 1973-1998, and include 104 and 312 observations for quarterly and monthly data, respectively. All exchanges rates are in natural logarithms and constructed using U.S. dollar as numeraire currency and CPI's as deflators.

Dealing with the presence of cross-sectional dependency has been one of the main econometric issues in the panel unit root testing. In the previous studies, it is commonly assumed away by imposing cross-sectional independence,[13] which can be quite unrealistic in many economic applications including the studies on the exchanges rates. Due to the strong links across economies, the real exchange rates usually exhibit high correlations across cross sections. Indeed this is seen clearly from the estimated correlation matrix of the first differences of the quarterly real exchange rates presented in Table 6. They are highly correlated especially among European countries. Pairwise correlations are particularly strong (over 0.9) among Austria, Belgium, Denmark, France, Germany, and the Netherlands. The pairwise correlations among the other countries are also significant and most of them are over 0.6. These numbers suggest that the overall cross-correlations among the real exchange rates are quite substantial and therefore should not be ignored.[14] Moreover, such high correlations among the exchange rates from European countries indicate that there may exist some cointegrating relations among them. Note that our tests allow for dependent and possibly cointegrated panels.

We carry out panel unit root testing for the full panels and also for the smaller panels which exclude the exchange rates from Australia, Canada, Greece, Japan, and Portugal. The smaller panels are considered to see if the observation made by Papell (2001) also holds in our study. Papell (2001) observes that the exchange rates from the five aforementioned countries follow different behavioral patterns, and that removal of these countries from the panel enables one to reject the unit root null hypothesis. The results are provided in Tables 7 and 8 for the quarterly and monthly data, respectively. In each table, the first line reports the results obtained from using the full panel, while the second line provides the results from using the smaller panel. The order of the lagged differences is selected for each cross section by the BIC criterion with the maximum lag order set at 12.[15] For both quarterly and monthly data, our tests, S^H and S^A, do not reject the unit root null hypothesis. This is the case regardless of whether we exclude those five countries from the full panel or not. That the tests S^H and S^A, constructed with and without the covariate, produced the same results in all cases indicates that the use of the covariate does not affect the results. We note in particular that all of the minimum tests, $S^C_{\min}$ and $S^F_{\min}$, $S^H_{\min}$ and

[12]The countries considered include Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.

[13]O'Connell (1998) controls for cross-sectional dependence by using the orthogonalized data which are obtained by the GLS transformation based on the estimated covariance matrix of the data.

[14]Similar observations were made for monthly real exchange rates in Moon and Perron (2001).

[15]The BIC selected 1 for the lag order for most of the cross sections. The results are not very sensitive to the lag order used.

$S^A_{\min}$, do not reject the null hypothesis for all cases considered. Recall from the simulation results reported in the previous section that all of the minimum tests show stable sizes for the sample sizes of our data, and have more discriminatory power over any average tests especially when only a fraction of the panel are stationary. Therefore, the results from the minimum tests strongly indicate that there is no stationary series in the panels of the real exchange rates.

In sharp contrast, the S^C, S^F, and IPS tests provide strong evidence against the unit root null hypothesis. When the full panels are used, the S^C test rejects the null for both quarterly and monthly data, and the S^F and IPS tests for quarterly data. If we exclude the five aforementioned countries from the panel, these results get strengthened and the S^F and IPS tests now reject the null also for the monthly data. The test results of S^C, S^F and IPS, however, may be spurious due to the likely existence of cointegrating relations among the cross sections in our data. These tests do not allow for cross-sectional cointegration and thus suffer from serious upward size distortions, as we saw from our simulation experiments with cointegrated panels. On the other hand, the MP test does not reject the null hypothesis in most of the cases considered. It rejects only for the quarterly data with the smaller panel.[16] However, these results may also be misleading, again due to the presence of cointegration, which the test is not designed to deal with. We also saw from our simulations that the MP test may have either upward or downward size distortions depending upon the time series and cross-sectional dimensions.

That our new average tests, S^H and S^A, and all the minimum tests, $S^C_{\min}$, $S^F_{\min}$, $S^H_{\min}$ and $S^A_{\min}$, do not provide any evidence in favor of the PPP hypothesis is in sharp contrast with most of the results in the previous literature. See Chang (2002) and Wu and Wu (2001) for some recent examples. All of the previous results were, however, obtained using the tests that assume cross-sectional independence and/or no cointegration. As our simulation experiments demonstrated, such tests suffer greatly from size distortions when there are strong cross-correlations induced by common stochastic trends and therefore are not suitable for cointegrated panels. On the other hand, our tests are designed to handle the cointegration among cross sections as well as cross-correlated regression errors, and hence our results against the PPP hypothesis seem reliable and appealing.

6. Conclusions

This paper extends the existing methodologies for panel unit root tests in three important directions. First, we allow for dependencies across individual cross sections at both short-run and long-run levels. We allow for inter-relatedness of cross-sectional short-run dynamics and the presence of long-run relationships in cross-sectional levels. Many panels of practical interest seem to have such complicated cross-sectional dependencies. Second, our theory permits the use of covariates to increase the power. Covariates may naturally include the terms to account for cross-sectional dependencies as well as the ones to control idiosyncrasies of individual cross-sectional units. If properly chosen, the inclusion

[16] These results are consistent with those of Moon and Perron (2001), which are obtained using the period-average real exchange rates for the period 1974-1998.

of covariates would substantially improve the power of the test, as demonstrated earlier by several authors. Third, we re-examine the formulation of the unit root hypothesis in panels, and propose to analyze the null and alternative hypotheses that only a fraction of cross-sectional units have unit roots. Such formulations are more appropriate for some of the most commonly investigated panel models such as purchasing power parity and growth convergence.

The tests developed in the paper are valid for very general panels. They allow not only for unknown forms of cross-sectional dependencies at several different levels, but also for various kinds of heterogeneities such as unbalancedness, differing dynamics and other idiosyncratic characteristics for individual units. These indeed appear to be the common characteristics of many panels used in empirical studies. Nevertheless, none of the currently available tests are applicable for such general panels. In addition to their applicability, our tests are easy to implement. The relevant statistical theories are quite straightforward and all Gaussian, and the critical values are given by either the standard normal or its simple functionals.

Appendix: Mathematical Proofs

Proof of Lemma 3.1 The asymptotics of the following sample moments involving integrable transformations of a unit root process follow directly from Park and Phillips (1999, 2001) as

$$\begin{aligned} T_i^{-1/4}\sum_{t=1}^{T_i} F_i(y_{i,t-1})\varepsilon_{it} &\to_d \mathbf{MN}\left(0, \sigma_i^2 L_i(1,0)\int_{-\infty}^{\infty} F_i(s)^2 ds\right), \\ T_i^{-1/2}\sum_{t=1}^{T_i} F_i(y_{i,t-1})^2 &\to_d L_i(1,0)\int_{-\infty}^{\infty} F_i(s)^2 ds. \end{aligned}$$

Note that our asymptotic results here are different, up to a scalar factor, from Chang (2002) that represents the asymptotics in terms of the local time of the standard Brownian motion. Since $\triangle y_{i,t-k}$, $k = 1, \ldots, P_i$, and $w_{i,t-j}$, $j = 1, \ldots, Q_i$, are stationary, we also have

$$\begin{aligned} T_i^{-3/4}\sum_{t=1}^{T_i} F_i(y_{i,t-1})\triangle y_{i,t-k} &\to_p 0, \quad \text{for all } k = 1, \ldots, P_i, \\ T_i^{-3/4}\sum_{t=1}^{T_i} F_i(y_{i,t-1}) w_{i,t-j} &\to_p 0, \quad \text{for all } j = 1, \ldots, Q_i, \end{aligned}$$

due to the asymptotic orthogonality between stationary variables and integrable transformations of integrated processes established in Lemma 5 (e) of Chang, Park and Phillips (2001).

Using (9) and (11), we may write τ_i defined in (10) as

$$\tau_i = \frac{B_{T_i}^{-1}A_{T_i}}{(\hat{\sigma}_i^2 B_{T_i}^{-2}C_{T_i})^{1/2}} = \frac{A_{T_i}}{\hat{\sigma}_i C_{T_i}^{1/2}} = \frac{T_i^{-1/4}\displaystyle\sum_{t=1}^{T_i} F_i(y_{i,t-1})\varepsilon_{it}}{\hat{\sigma}_i\left(T_i^{-1/2}\displaystyle\sum_{t=1}^{T_i} F_i(y_{i,t-1})^2\right)^{1/2}} \; + \; o_p(1).$$

Now the stated result follows immediately. □

Proof of Lemma 3.2 The asymptotic independence of $\tau_1, \ldots, \tau_M$ follows if we show that τ_i and τ_j are asymptotically orthogonal for all $i, j = 1, \ldots, M$. The proof goes exactly the same as that in Chang (2002), except for the pairs τ_i and τ_j for which the corresponding cross-sectional units (y_{it}) and (y_{jt}) are cointegrated. Note that Assumption 4.1 in Chang (2002) holds under our Assumption 3.2. Therefore, we assume (y_{it}) and (y_{jt}) are cointegrated

To establish the asymptotic orthogonality of τ_i and τ_j, it suffices to show that

$$\sqrt[4]{T_i T_j}\int_0^1 F_i(\sqrt{T_i}U_{iT_i}(r))F_j(\sqrt{T_j}U_{jT_j}(r))dr \to_p 0. \tag{26}$$

See Chang (2002) for details. As shown in Chang, Park and Phillips (2001), we have

$$\begin{aligned}&\frac{T_1}{\log T_1}\int_0^1 F_i(\sqrt{T_1}U_{iT_i}(r))F_j(\sqrt{T_1}U_{jT_j}(r))dr \\ &= \frac{T_1}{\log T_1}\int_0^1 F_i(\sqrt{T_1}U_i(r))F_j(\sqrt{T_1}U_j(r))dr + o_p(1).\end{aligned}$$

However, due to our convention in (6) made on scale adjustment, we may assume that

$$\begin{aligned}&\int_0^1 F_i(\sqrt{T_i}U_{iT_i}(r))F_j(\sqrt{T_j}U_{jT_j}(r))dr \\ &= \int_0^1 F_i(\sqrt{T_1}U_{iT_i}(r))F_j(\sqrt{T_1}U_{jT_j}(r))dr\end{aligned}$$

and that

$$U_i = U_j.$$

Moreover, we have

$$\int_0^1 (F_iF_j)(\sqrt{T_1}U_i(r))dr = O_p(T_1^{-3/4}),$$

since F_i and F_j are assumed to be orthogonal and $\int_{-\infty}^{\infty}(F_iF_j)(s)ds = 0$. This can be deduced from

$$T_1^{-1/4}\int_0^{T_1}(F_iF_j)(U_i(r))dr = O_p(1),$$

which is shown in, e.g., Revuz and Yor (1994, Proposition 2.8, p.528). Notice that

$$T_1^{-1/4}\int_0^{T_1}(F_iF_j)(U_i(r))dr =_d T_1^{3/4}\int_0^1 (F_iF_j)(\sqrt{T_1}U_i(r))dr,$$

which follows immediately from the change-of-variable formula and the fact that $U_i(T_1 r) =_d \sqrt{T_1}U_i(r)$ for every $r \geq 0$.

We may assume without loss of generality that the numeraire unit has the largest number of observations, i.e., $T_1 \geq T_i$ for all i. Then we have

$$\begin{aligned}
&\sqrt[4]{T_iT_j}\int_0^1 F_i(\sqrt{T_i}U_{iT_i}(r))F_j(\sqrt{T_j}U_{jT_j}(r))dr \\
&= \sqrt[4]{T_iT_j}\int_0^1 F_i(\sqrt{T_1}U_{iT_i}(r))F_j(\sqrt{T_1}U_{jT_j}(r))dr \\
&= \sqrt[4]{T_iT_j}\int_0^1 (F_iF_j)(\sqrt{T_1}U_i(r))dr + o_p\left(\frac{T_i^{1/4}T_j^{1/4}\log T_1}{T_1}\right) \\
&= O_p\left(T_j^{1/4}/T_i^{1/2}\right) \\
&= O_p\left(T_{\max}^{1/4}/T_{\min}^{1/2}\right)
\end{aligned}$$

since T_1 is set at $T_{\max}$. Now (26) follows immediately given Assumption 3.2. □

Proof of Theorem 3.3 We first consider the case $M = N$. The statistic S has standard normal limiting distribution and, therefore, the stated result follows immediately. For the statistic $S_{\min}$, we note that

$$\begin{aligned}
\lim_{T\to\infty}\mathbf{P}\{S_{\min}\leq x\} &= \lim_{T\to\infty}\mathbf{P}\left\{\min_{1\leq i\leq N}\tau_i \leq x\right\} \\
&= 1-\prod_{i=1}^N \lim_{T_i\to\infty}\mathbf{P}\{\tau_i > x\} \\
&= 1-(1-\Phi(x))^N
\end{aligned}$$

since τ_i's are asymptotically independent normals.

For the case $1 \leq M \leq N$, we have

$$\begin{aligned}
\lim_{T\to\infty}\mathbf{P}\{S_{\max}\leq x\} &= \lim_{T\to\infty}\mathbf{P}\left\{\max_{1\leq i\leq N}\tau_i \leq x\right\} \\
&= \prod_{i=1}^N \lim_{T_i\to\infty}\mathbf{P}\{\tau_i \leq x\} \\
&= \Phi(x)^M.
\end{aligned}$$

Note that for $i = 1, \ldots, M$,

$$\lim_{T_i\to\infty}\mathbf{P}\{\tau_i \leq x\} = \Phi(x)$$

and for $i = M+1, \ldots, N$,

$$\lim_{T_i \to \infty} \mathbf{P}\{\tau_i \leq x\} = 1$$

since $\tau_i \to_p -\infty$ as $T_i \to \infty$ in this case. Therefore,

$$\lim_{T \to \infty} \mathbf{P}\left\{S_{\max} \leq c^M_{\max}(\lambda)\right\} = \Phi(c^M_{\max}(\lambda))^M = \lambda$$

and

$$\lim_{T \to \infty} \mathbf{P}\left\{S_{\max} \leq c(\lambda)\right\} = \Phi(c(\lambda))^M = \lambda \Phi(c(\lambda))^{M-1} \leq \lambda$$

as was to be shown. The consistency of the tests then follows immediately from the result in Lemma 3.1. □

Proof of Corollary 3.4 The proof of this corollary is essentially the same as that of Corollary 5.1 in Chang (2002). What is new is how to handle the covariates with deterministic components. Assume for simplicity that the panels are balanced, i.e., $T_i = T$ for all i. Note that

$$\begin{aligned}
\triangle y_{j,t-k} &= \triangle z_{j,t-k} = \triangle y_{j,t-k}, \quad \text{for } k = 1, 2, \ldots, Q_i, \\
(y_{i,t-1} - \phi_i y_{j,t-1}) &= \left(z_{i,t-1} - \frac{1}{T}\sum_{k=1}^T z_{ik}\right) - \phi_i \left(z_{j,t-1} - \frac{1}{T}\sum_{k=1}^T z_{jk}\right) \\
&= \left(y_{i,t-1} - \frac{1}{T}\sum_{k=1}^T y_{ik}\right) - \phi_i \left(y_{j,t-1} - \frac{1}{T}\sum_{k=1}^T y_{jk}\right) \\
&= (y_{i,t-1} - \phi_i y_{j,t-1}) - \frac{1}{T}\sum_{k=1}^T (y_{ik} - \phi_i y_{jk}), \qquad (27) \\
y_{jt} &= z_{jt} - \frac{1}{T}\sum_{k=1}^T z_{jk} = y_{jt} - \frac{1}{T}\sum_{k=1}^T y_{jk}. \qquad (28)
\end{aligned}$$

Define the cointegration errors as $\epsilon_{it} = y_{it} - \phi_i y_{jt}$, and let

$$v_{iT} = \frac{1}{T}\sum_{k=1}^T \epsilon_{ik} \quad \text{or} \quad \frac{1}{T}\sum_{k=1}^T y_{jk}, \qquad (29)$$

depending upon whether y_{jt} is cointegrated with y_{it} as in (27) or it is stationary as in (28), respectively. Then, it is sufficient to show that

$$R_{iT} = \frac{1}{\sqrt[4]{T}} \sum_{t=1}^T F_i(y^\mu_{i,t-1}) v_{iT}$$

is negligible in the limit and dominated by the leading term

$$\frac{1}{\sqrt[4]{T}} \sum_{t=1}^T F_i(y^\mu_{i,t-1}) \varepsilon_{it},$$

so that the regression equation (15) is the legitimate equation for the test of unit roots in (y_{it}). This follows immediately from

$$R_{iT} = \frac{1}{\sqrt[4]{T}}\left(\sqrt{T}v_{iT}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} F_i(y_{i,t-1}^{\mu})\right) = O_p(T^{-1/4}), \tag{30}$$

by Lemma 5 (a) in Chang, Park, and Phillips (2001).

For the detrended series, first consider the term $\frac{1}{T}z_{jT} = \frac{1}{T}\sum_{k=1}^{T}\triangle z_{jk}$ in (21). We have $y_{it} = y_{i,t-1} + u_{it}$ under the unit root null, and this implies $\triangle z_{it} = \delta_i + \triangle y_{it} = \delta_i + u_{it}$. Then, we have

$$\begin{aligned}\triangle y_{j,t-k} &= \triangle z_{j,t-k} - \frac{1}{T}z_{jT_j} \\ &= u_{j,t-k} - \frac{1}{T}\sum_{k=1}^{T} u_{jk}.\end{aligned}$$

By letting $v_{iT} = \frac{1}{T}\sum_{k=1}^{T} u_{jk}$, the second term can be handled in the same manner as in (30). For $y_{i,t-1}^{\tau}$ in (22), notice that it can be written as follows:

$$\begin{aligned}y_{i,t-1}^{\tau} &= z_{i,t-1} - \hat{\mu}_i^T - \hat{\delta}_i^T(t-1) \\ &= y_{i,t-1} - \frac{1}{T}\sum_{k=1}^{T} y_{ik} + \frac{6}{T(T-1)}\sum_{k=1}^{T}\left(k - \frac{T+1}{2}\right)y_{ik} \\ &\quad -(t-1)\left[\frac{12}{T(T-1)(T+1)}\sum_{k=1}^{T}\left(k - \frac{T+1}{2}\right)y_{ik}\right],\end{aligned} \tag{31}$$

where $\hat{\mu}_i^T$ and $\hat{\delta}_i^T$ are the LS estimators for the parameters μ_i and δ_i in (19) using the full sample, viz.,

$$\begin{aligned}\hat{\mu}_i^T + \frac{T+1}{2}\hat{\delta}_i^T &= \mu_i + \frac{T+1}{2}\delta_i + \frac{1}{T}\sum_{k=1}^{T} y_{ik}, \\ \hat{\delta}_i^T &= \delta_i + \left(\sum_{k=1}^{T}\left(k - \frac{T+1}{2}\right)^2\right)^{-1}\sum_{k=1}^{T}\left(k - \frac{T+1}{2}\right)y_{ik}.\end{aligned}$$

Now, along with $y_{j,t-1}^{\tau}$ defined in the same way, the first term in (31) forms the lagged cointegration errors, $(y_{i,t-1} - \phi_i y_{j,t-1})$, and is used as covariates. The remaining terms can be shown to be negligible in the limit as in (30) with slight rearrangement of the terms. For the second term, we may use the relation $\epsilon_{it} = y_{i,t} - \phi_i y_{j,t}$ to generate v_{iT} as in (29). The third term can be rewritten as

$$v_{iT} = \frac{6}{T-1}\sum_{k=1}^{T}\left(\frac{k}{T}\right)\epsilon_{ik} - \frac{3(T+1)}{T(T-1)}\sum_{k=1}^{T}\epsilon_{ik}.$$

For the fourth term, we can show that

$$R_{iT} = \frac{1}{\sqrt[4]{T}} \left(\sqrt{T} v_{iT} \right) \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_i(y^{\tau}_{i,t-1}) \left(\frac{t-1}{T} \right) \right) = O_p(T^{-1/4}),$$

where

$$v_{iT} = \frac{12T}{(T-1)(T+1)} \sum_{k=1}^{T} \left(\frac{k}{T} \right) \epsilon_{ik} - \frac{6}{T-1} \sum_{k=1}^{T} \epsilon_{ik},$$

by Lemma 5 (g) in Chang, Park, and Phillips (2001). The case of the stationary covariate y^{τ}_{jt} can be handled in the similar manner as above, and thus it is omitted. □

References

Bai, J., and S. Ng (2002). "Determining the number of factors in approximate factor models," *Econometrica*, 70, 191-221.

Baltagi, B.H., and C. Kao (2000). "Nonstationary panels, cointegration in panels and dynamic panels: A survey," mimeographed.

Banerjee, A. (1999). "Panel data unit roots and cointegration: An overview," *Oxford Bulletin of Economics and Statistics*, 61, 607-629.

Chang, Y. (2002). "Nonlinear IV unit root tests in panels with cross-sectional dependency," *Journal of Econometrics*, 110, 261-292.

Chang, Y. (2004). "Bootstrap unit root tests in panels with cross-sectional dependency," *Journal of Econometrics*, 120, 263-293.

Chang, Y., J.Y. Park and P.C.B. Phillips (2001). "Nonlinear econometric models with cointegrated and deterministically trending regressors," *Econometrics Journal*, 4, 1-36.

Chang, Y., R. Sickles, and W. Song (2001), "Bootstrapping unit root tests with covariates," Rice University, mimeographed.

Choi, I. (2001a). "Unit root tests for panel data," *Journal of International Money and Finance*, 20, 219-247.

Choi, I. (2001b). "Unit root tests for cross-sectionally correlated panels," Kookmin University, mimeographed.

Hansen, B.E. (1995). "Rethinking the univariate approach to unit root testing: Using covariates to increase power," *Econometric Theory*, 11, 1148-1171.

Im, K., H. Pesaran and Y. Shin (2003). "Testing for unit roots in heterogeneous panels," *Journal of Econometrics* 115, pp.53-74.

Levin, A., C. Lin and C.J. Chu (2002). "Unit root tests in panel data: Asymptotic and finite-sample properties," *Journal of Econometrics* 108, 1-24.

MacDonald, R. (1996). "Panel unit root tests and real exchange rates," *Economics Letters*, 50, 7-11.

Maddala, G.S. and S. Wu (1999). "A comparative study of unit root tests with panel data and a new simple test," *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

Moon, H.R. and B. Perron (2001). "Testing for a unit root in panels with dynamic factors," mimeographed.

O'Connell, P.G.J. (1998). "The overvaluation of purchasing power parity," *Journal of International Economics*, 44, 1-19.

Oh, K.Y. (1996). "Purchasing power parity and unit root tests using panel data," *Journal of International Money and Finance*, 15, 405-418.

Papell, D.H. (1997). "Searching for stationarity: Purchasing power parity under the current float," *Journal of International Economics*, 43, 313-323.

Papell, D.H. (2000). "The purchasing power parity persistence paradigm," forthcoming in *Journal of International Economics.*

Papell, D.H. (2001). "The great appreciation, the great depreciation, and the purchasing power parity hypothesis," forthcoming in *Journal of International Economics.*

Park, J.Y. and P.C.B. Phillips (1999). "Asymptotics for nonlinear transformations of integrated time series," *Econometric Theory*, 15, 269-298.

Park, J.Y. and P.C.B. Phillips (2001). "Nonlinear regressions with integrated time series," *Econometrica*, 69, 117-162.

Phillips, P.C.B. and H.R. Moon (2000). "Nonstationary panel data analysis: An overview of some recent developments," *Econometric Reviews,* 19, 263-86.

Phillips, P.C.B. and D. Sul (2001). "Dynamic panel estimation and homogeneity testing under cross section dependence," mimeographed.

Revuz, D. and M. Yor (1994). *Continuous Martingale and Brownian Motion,* New Your: Springer-Verlag.

Wu, J.L. and S. Wu (2001). "Is purchasing power parity overvalued?," *Journal of Money, Credit and Banking*, 33, 804-12.

Table 1: Critical Values for $S_{\max}$ and $S_{\min}$

$S_{\max}$				$S_{\min}$			
M	1%	5%	10%	N	1%	5%	10%
2	-1.282	-0.760	-0.478	2	-2.575	-1.955	-1.632
5	-0.258	0.124	0.334	5	-2.877	-2.319	-2.036
10	0.334	0.647	0.822	10	-3.089	-2.568	-2.309
13	0.529	0.821	0.985	13	-3.166	-2.657	-2.406
15	0.630	0.911	1.070	15	-3.207	-2.705	-2.457
17	0.715	0.988	1.142	17	-3.243	-2.746	-2.502
20	0.822	1.084	1.233	20	-3.289	-2.799	-2.559
25	0.961	1.211	1.353	25	-3.351	-2.870	-2.635
30	1.070	1.310	1.447	30	-3.402	-2.928	-2.696
40	1.233	1.460	1.590	40	-3.479	-3.016	-2.791
50	1.353	1.570	1.695	50	-3.539	-3.083	-2.862
60	1.447	1.658	1.779	60	-3.587	-3.137	-2.919
70	1.525	1.729	1.847	70	-3.627	-3.182	-2.967
80	1.590	1.790	1.905	80	-3.661	-3.220	-3.008
90	1.646	1.842	1.956	90	-3.691	-3.254	-3.043
100	1.695	1.888	2.000	100	-3.718	-3.283	-3.075

Table 2

Sizes and Size-Adjusted Powers for DGP1: Hypotheses (A) and (B)

N	T	I_0	S^C	$S^C_{\min}$	S^F	$S^F_{\min}$	S^H	$S^H_{\min}$	S^A	$S^A_{\min}$	IPS	MP
10	100	0	0.071	0.064	0.062	0.063	0.024	0.052	0.043	0.073	0.080	0.104
		1	0.085	0.102	0.085	0.101	0.110	0.131	0.110	0.113	0.073	0.049
		2	0.181	0.245	0.175	0.242	0.261	0.329	0.266	0.285	0.141	0.047
		5	0.671	0.524	0.627	0.476	0.501	0.613	0.679	0.531	0.497	0.053
		10	0.997	0.724	0.994	0.653	0.609	0.733	0.851	0.682	0.961	0.182
	200	0	0.071	0.062	0.062	0.059	0.032	0.052	0.051	0.068	0.072	0.072
		1	0.158	0.472	0.161	0.462	0.255	0.550	0.229	0.514	0.151	0.051
		2	0.334	0.652	0.323	0.639	0.498	0.751	0.482	0.703	0.284	0.047
		5	0.937	0.951	0.919	0.930	0.896	0.977	0.959	0.954	0.872	0.050
		10	1.000	0.994	1.000	0.989	0.969	0.996	0.998	0.993	1.000	0.119
20	100	0	0.059	0.067	0.055	0.065	0.012	0.034	0.029	0.060	0.075	0.146
		2	0.160	0.197	0.159	0.198	0.254	0.347	0.263	0.263	0.127	0.047
		5	0.472	0.378	0.455	0.377	0.424	0.586	0.595	0.485	0.347	0.048
		10	0.939	0.587	0.925	0.562	0.531	0.724	0.806	0.656	0.806	0.059
		20	1.000	0.807	1.000	0.773	0.661	0.789	0.912	0.795	1.000	0.337
	200	0	0.056	0.065	0.052	0.064	0.020	0.033	0.042	0.059	0.061	0.082
		2	0.230	0.526	0.230	0.525	0.361	0.689	0.344	0.618	0.193	0.050
		5	0.767	0.821	0.754	0.812	0.760	0.914	0.866	0.874	0.657	0.051
		10	0.999	0.984	0.999	0.979	0.909	0.990	0.991	0.991	0.994	0.065
		20	1.000	1.000	1.000	1.000	0.966	0.999	1.000	1.000	1.000	0.579
50	100	0	0.055	0.071	0.053	0.071	0.005	0.016	0.014	0.033	0.080	0.379
		5	0.228	0.204	0.227	0.201	0.230	0.491	0.343	0.373	0.174	0.035
		10	0.578	0.339	0.570	0.330	0.291	0.647	0.528	0.564	0.405	0.028
		25	0.999	0.634	0.999	0.615	0.431	0.738	0.725	0.780	0.981	0.027
		50	1.000	0.855	1.000	0.833	0.639	0.741	0.865	0.820	1.000	0.231
	200	0	0.052	0.070	0.051	0.069	0.007	0.015	0.021	0.033	0.064	0.165
		5	0.449	0.841	0.443	0.827	0.640	0.985	0.722	0.958	0.375	0.048
		10	0.906	0.963	0.902	0.957	0.803	0.998	0.959	0.996	0.815	0.044
		25	1.000	1.000	1.000	1.000	0.924	1.000	0.999	1.000	1.000	0.050
		50	1.000	1.000	1.000	1.000	0.978	1.000	1.000	1.000	1.000	0.622

Note: I_0 denotes the number of I(0) series in the panel under the alternative hypothesis of Hypotheses (B). The sizes are reported in the rows corresponding to $I_0 = 0$, and the size-adjusted powers in the rows with $I_0 > 0$.

Table 3

Sizes and Size-Adjusted Powers for DGP2: Hypotheses (A) and (B)

		I_0	S^C	$S^C_{\min}$	S^F	$S^F_{\min}$	S^H	$S^H_{\min}$	S^A	$S^A_{\min}$	IPS	MP
10	100	0	0.216	0.058	0.037	0.057	0.038	0.048	0.030	0.058	0.229	0.088
		1	0.066	0.099	0.069	0.290	0.103	0.131	0.149	0.342	0.056	0.047
		2	0.103	0.238	0.190	0.764	0.230	0.329	0.474	0.823	0.076	0.054
		5	0.307	0.434	0.591	0.884	0.427	0.510	0.802	0.920	0.200	0.101
		10	0.810	0.565	0.910	0.716	0.506	0.580	0.718	0.752	0.587	0.260
	200	0	0.210	0.055	0.038	0.051	0.046	0.046	0.038	0.055	0.217	0.062
		1	0.102	0.478	0.213	0.781	0.225	0.563	0.444	0.820	0.084	0.052
		2	0.179	0.658	0.492	0.942	0.445	0.754	0.804	0.958	0.138	0.063
		5	0.606	0.915	0.795	0.997	0.813	0.954	0.950	0.999	0.496	0.099
		10	0.973	0.973	0.980	0.985	0.891	0.979	0.977	0.992	0.927	0.142
20	100	0	0.288	0.056	0.046	0.059	0.024	0.030	0.023	0.044	0.315	0.110
		2	0.082	0.189	0.189	0.726	0.218	0.326	0.461	0.830	0.066	0.062
		5	0.161	0.347	0.414	0.911	0.361	0.516	0.779	0.958	0.112	0.095
		10	0.374	0.467	0.737	0.961	0.433	0.587	0.861	0.980	0.239	0.148
		20	0.859	0.614	0.978	0.855	0.542	0.623	0.706	0.870	0.649	0.384
	200	0	0.292	0.055	0.046	0.055	0.033	0.031	0.030	0.043	0.303	0.069
		2	0.098	0.520	0.276	0.884	0.311	0.676	0.649	0.916	0.084	0.069
		5	0.264	0.769	0.623	0.981	0.676	0.872	0.958	0.992	0.204	0.109
		10	0.683	0.943	0.963	1.000	0.807	0.966	0.998	1.000	0.559	0.166
		20	0.998	0.987	0.999	0.996	0.878	0.978	0.967	0.998	0.987	0.471
50	100	0	0.374	0.056	0.052	0.062	0.020	0.014	0.009	0.023	0.407	0.183
		5	0.080	0.189	0.169	0.871	0.190	0.427	0.630	0.959	0.064	0.080
		10	0.122	0.290	0.392	0.936	0.231	0.537	0.772	0.982	0.084	0.115
		25	0.358	0.473	0.857	0.986	0.324	0.580	0.798	0.993	0.223	0.200
		50	0.844	0.607	0.995	0.930	0.481	0.588	0.584	0.896	0.619	0.421
	200	0	0.378	0.052	0.053	0.060	0.019	0.014	0.017	0.023	0.401	0.102
		5	0.116	0.776	0.488	0.999	0.516	0.959	0.971	1.000	0.093	0.107
		10	0.216	0.880	0.677	1.000	0.653	0.976	0.997	1.000	0.163	0.141
		25	0.711	0.978	0.995	1.000	0.776	0.988	0.999	1.000	0.583	0.221
		50	0.998	0.995	1.000	1.000	0.890	0.989	0.986	1.000	0.985	0.456

Note: I_0 denotes the number of I(0) series in the panel under the alternative hypothesis of Hypotheses (B). The sizes are reported in the rows corresponding to $I_0 = 0$, and the size-adjusted powers in the rows with $I_0 > 0$.

Table 4

Sizes and Size-Adjusted Powers for DGP3: Hypotheses (A) and (B)

		I_0	S^C	$S^C_{\min}$	S^F	$S^F_{\min}$	S^H	$S^H_{\min}$	S^A	$S^A_{\min}$	IPS	MP
10	100	0	0.330	0.048	0.180	0.020	0.063	0.038	0.054	0.021	0.329	0.250
		1	0.052	0.138	0.051	0.400	0.074	0.192	0.121	0.456	0.047	0.027
		2	0.063	0.300	0.110	0.901	0.169	0.417	0.446	0.927	0.053	0.016
		5	0.115	0.501	0.374	0.951	0.354	0.592	0.758	0.969	0.084	0.012
		10	0.418	0.642	0.684	0.885	0.417	0.656	0.644	0.914	0.286	0.124
	200	0	0.326	0.044	0.180	0.017	0.073	0.038	0.061	0.018	0.320	0.043
		1	0.066	0.507	0.145	0.890	0.135	0.606	0.404	0.911	0.065	0.192
		2	0.085	0.709	0.362	0.987	0.336	0.800	0.812	0.996	0.079	0.227
		5	0.241	0.949	0.660	1.000	0.737	0.979	0.933	1.000	0.217	0.270
		10	0.818	0.984	0.873	0.998	0.825	0.989	0.962	0.999	0.728	0.349
20	100	0	0.375	0.046	0.263	0.021	0.053	0.033	0.039	0.021	0.383	0.529
		2	0.058	0.257	0.093	0.903	0.162	0.411	0.456	0.950	0.053	0.017
		5	0.078	0.444	0.213	0.984	0.311	0.601	0.783	0.995	0.064	0.010
		10	0.124	0.524	0.397	0.991	0.366	0.633	0.837	0.996	0.087	0.008
		20	0.449	0.643	0.662	0.947	0.443	0.652	0.623	0.935	0.296	0.061
	200	0	0.377	0.044	0.279	0.020	0.057	0.024	0.042	0.018	0.378	0.263
		2	0.063	0.553	0.155	0.914	0.220	0.717	0.633	0.946	0.058	0.013
		5	0.106	0.793	0.367	0.997	0.590	0.883	0.956	0.999	0.092	0.007
		10	0.255	0.969	0.752	1.000	0.716	0.980	0.997	1.000	0.213	0.005
		20	0.897	0.992	0.951	1.000	0.785	0.987	0.932	1.000	0.808	0.171
50	100	0	0.422	0.047	0.367	0.023	0.051	0.016	0.018	0.012	0.438	0.739
		5	0.054	0.238	0.064	0.944	0.150	0.505	0.648	0.981	0.053	0.012
		10	0.068	0.394	0.102	0.976	0.188	0.632	0.784	0.996	0.055	0.008
		25	0.115	0.581	0.388	0.995	0.245	0.661	0.766	0.996	0.081	0.006
		50	0.382	0.683	0.604	0.964	0.350	0.667	0.488	0.938	0.252	0.026
	200	0	0.428	0.046	0.415	0.021	0.048	0.015	0.026	0.013	0.437	0.607
		5	0.066	0.786	0.150	1.000	0.396	0.965	0.981	1.000	0.065	0.005
		10	0.085	0.868	0.232	1.000	0.532	0.983	0.997	1.000	0.079	0.003
		25	0.241	0.977	0.778	1.000	0.651	0.992	0.998	1.000	0.215	0.001
		50	0.870	0.993	0.990	1.000	0.772	0.992	0.968	1.000	0.781	0.022

Note: I_0 denotes the number of I(0) series in the panel under the alternative hypothesis of Hypotheses (B). The sizes are reported in the rows corresponding to $I_0 = 0$, and the size-adjusted powers in the rows with $I_0 > 0$.

Table 5. Sizes and Size-Adjusted Powers for DGP1 and DGP2: Hypotheses (C)

			DGP1				DGP2			
			Sizes		Powers		Sizes		Powers	
N	T	I_1	$S^C_{\max}$	$S^F_{\max}$	$S^C_{\max}$	$S^F_{\max}$	$S^C_{\max}$	$S^F_{\max}$	$S^C_{\max}$	$S^F_{\max}$
10	100	1	0.000	0.000	0.114	0.113	0.003	0.000	0.109	0.105
		2	0.007	0.003	0.225	0.219	0.033	0.000	0.188	0.187
		5	0.040	0.028	0.529	0.486	0.137	0.001	0.364	0.270
		10	0.054	0.048	0.762	0.720	0.213	0.040	0.558	0.158
	200	1	0.001	0.000	0.190	0.186	0.007	0.000	0.154	0.155
		2	0.015	0.007	0.349	0.329	0.051	0.000	0.277	0.211
		5	0.043	0.030	0.657	0.621	0.145	0.002	0.502	0.145
		10	0.055	0.051	0.822	0.785	0.215	0.044	0.659	0.083
20	100	2	0.001	0.000	0.165	0.163	0.016	0.000	0.137	0.147
		5	0.018	0.013	0.374	0.378	0.097	0.000	0.262	0.268
		10	0.033	0.029	0.637	0.623	0.188	0.000	0.417	0.248
		20	0.038	0.039	0.830	0.812	0.264	0.047	0.586	0.135
	200	2	0.012	0.007	0.374	0.374	0.055	0.000	0.292	0.264
		5	0.037	0.027	0.737	0.732	0.128	0.000	0.591	0.346
		10	0.042	0.037	0.903	0.897	0.201	0.000	0.755	0.323
		20	0.038	0.039	0.971	0.967	0.266	0.046	0.869	0.132
50	100	5	0.001	0.001	0.203	0.206	0.050	0.000	0.131	0.124
		10	0.011	0.010	0.385	0.383	0.150	0.000	0.210	0.162
		25	0.026	0.026	0.697	0.694	0.278	0.000	0.397	0.157
		50	0.037	0.038	0.849	0.844	0.373	0.061	0.558	0.053
	200	5	0.008	0.007	0.355	0.348	0.106	0.000	0.207	0.167
		10	0.023	0.021	0.593	0.595	0.182	0.000	0.372	0.143
		25	0.031	0.031	0.844	0.840	0.286	0.000	0.603	0.124
		50	0.030	0.031	0.933	0.933	0.342	0.048	0.757	0.037

Note: I_1 denotes the number of I(1) series in the panel under the null hypothesis of Hypotheses (C). The size-adjusted rejection probabilities are reported as powers.

Table 6. Correlation Matrix of First Differenced Quarterly Log Real Exchange Rates

	Aust.	Belg.	Denm.	Finl.	Fran.	Germ.	Gree.	Irel.	Ital.	Neth.	Norw.	Port.	Spai.	Swed.	Swit.	U.K.	Aust.	Cana.	Japa.
Austria																			
Belgium	.965																		
Denmark	.955	.966																	
Finland	.749	.750	.751																
France	.926	.927	.918	.723															
Germany	.982	.962	.955	.733	.925														
Greece	.674	.682	.715	.601	.718	.699													
Ireland	.845	.860	.864	.743	.863	.845	.727												
Italy	.733	.744	.749	.691	.803	.729	.658	.761											
Netherl.	.980	.970	.961	.752	.930	.980	.709	.863	.765										
Norway	.876	.854	.847	.777	.843	.871	.663	.802	.669	.852									
Portugal	.807	.807	.795	.676	.788	.809	.617	.717	.655	.802	.771								
Spain	.733	.725	.734	.687	.757	.708	.649	.726	.753	.744	.706	.681							
Sweden	.731	.733	.733	.788	.708	.723	.527	.662	.681	.726	.802	.721	.698						
Switzer.	.877	.859	.863	.684	.858	.875	.659	.768	.693	.871	.764	.711	.620	.659					
U.K.	.604	.626	.621	.714	.662	.619	.650	.768	.651	.640	.677	.593	.630	.671	.581				
Australia	.223	.205	.213	.244	.225	.199	.268	.218	.201	.207	.264	.155	.194	.193	.218	.224			
Canada	−.007	−.020	.026	.088	−.052	−.023	.028	−.014	−.037	−.033	.018	.063	−.069	.029	.023	.030	.360		
Japan	.605	.593	.608	.438	.574	.601	.517	.518	.490	.606	.508	.456	.440	.399	.633	.464	.278	.065	
New Zeal.	.440	.408	.417	.414	.407	.422	.464	.427	.359	.425	.450	.314	.318	.331	.390	.422	.632	.185	.422

Table 7. Test Results for Quarterly Real Exchange Rates

N	T	S^C	$S^C_{\min}$	S^F	$S^F_{\min}$	S^H	$S^H_{\min}$	S^A	$S^A_{\min}$	IPS	MP
20	104	-3.792*	-1.925	-3.054*	-2.387	-0.185	-1.452	-1.411	-2.169	-2.095*	-1.134
15	104	-4.057*	-1.925	-2.200*	-2.287	-0.806	-1.467	-0.935	-2.517	-2.794*	-2.012*

Notes: 1. The results in the second line are obtained excluding Australia, Canada, Greece, Japan, and Portugal.
2. * indicates significance at the 5% level.

Table 8. Test Results for Monthly Real Exchange Rates

N	T	S^C	$S^C_{\min}$	S^F	$S^F_{\min}$	S^H	$S^H_{\min}$	S^A	$S^A_{\min}$	IPS	MP
17	312	-2.576*	-1.277	-1.390	-1.642	-0.315	-1.100	-0.489	-1.651	-1.271	-0.328
13	312	-2.817*	-1.277	-2.171*	-1.737	0.407	-1.151	-1.094	-1.877	-1.720*	-0.375

Notes: 1. The results in the second line are obtained excluding Canada, Greece, Japan, and Portugal.
2. * indicates significance at the 5% level.